

**Instituto Politécnico de Coimbra**

Instituto Superior de Contabilidade  
e Administração de Coimbra

Vanessa Sofia Neves Abreu e Sousa

## **O Poder dos Dados no Jornalismo Digital: o caso do jornal Expresso**

O Poder dos dados no Jornalismo Digital: o caso do jornal Expresso

Vanessa Sousa

ISCAC | 2018

Coimbra, outubro de 2018





**Instituto Politécnico de Coimbra**

Instituto Superior de Contabilidade  
e Administração de Coimbra

Vanessa Sofia Neves Abreu e Sousa

## **O Poder dos Dados no Jornalismo Digital: o caso do jornal Expresso**

Relatório de Estágio submetido ao Instituto Superior de Contabilidade e Administração de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Sistemas de Informação de Gestão**, realizado sob a orientação da Professora Isabel Pedrosa, coorientação da Professora Sara Meireles e supervisão de Germano Oliveira.

Coimbra, outubro de 2018

## **TERMO DE RESPONSABILIDADE**

Declaro ser a autora deste relatório de estágio, que constitui um trabalho original e inédito, que nunca foi submetido a outra Instituição de ensino superior para obtenção de um grau acadêmico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas e que tenho consciência de que o plágio constitui uma grave falta de ética, que poderá resultar na anulação do presente relatório de estágio.



## **AGRADECIMENTOS**

A todos os que me abriram portas e me deram a possibilidade de fazer mais e melhor, o meu muito obrigada. Ao João Pereira, que me abriu horizontes. À professora Isabel Pedrosa e à professora Sara Meireles que viram potencial em mim e que de perto acompanharam e orientaram o desenvolvimento do presente relatório. Ao Germano Oliveira, editor executivo do jornal Expresso, que me permitiu fazer parte da redação e sempre me motivou. Ao editor adjunto José Cardoso e ao coordenador multimédia Paulo Luís de Castro pelo acompanhamento diário. À jornalista Raquel Albuquerque por todo o apoio e partilha de conhecimentos. À jornalista Carolina Reis por toda a motivação e pelas experiências únicas que me proporcionou. À jornalista Manuela Goucha Soares pela confiança e reconhecimento. À coordenadora do 2:59 Joana Beleza, ao realizador André Atayde e ao editor multimédia Carlos Paes por me integrarem na rubrica. A todos os jornalistas que acompanharam o meu trabalho. À secretária do diretor Ana Bela por estar sempre disposta a ajudar. Ao senhor Abreu e à dona Celeste por sempre acreditarem em mim, não esquecendo da minha avó Ilda, da minha madrinha Olga, da minha irmã Patrícia, do meu sobrinho Ângelo e do meu primo Alexandre. Muito obrigada por tudo, a todos.



## RESUMO

Se há poder que o jornalismo tem é o de se moldar às circunstâncias, às adversidades, aos avanços tecnológicos, a novos meios de transmissão de informação. Da sua adaptação à Internet nasce o jornalismo digital, que, por um lado, traz uma série de possibilidades, mas, por outro, cria um ambiente de instabilidade que abala a credibilidade do jornalismo. E da necessidade de compreender os biliões de dados que todos os dias são gerados no meio digital floresce o jornalismo de dados. Esta prática, que extrai histórias de uma avalanche de dados, tornou-se numa nova tendência jornalística internacional. Já são vários os *media* a integrar o jornalismo de dados nas suas redações. E o jornal Expresso não é exceção. Tendo em vista a compreensão deste fenómeno, os seis meses de estágio no Expresso Digital, as leituras, as entrevistas e os questionários – compilados no presente relatório – procuram explorar as mudanças que a integração do jornalismo de dados gera nesta redação, nas rotinas produtivas dos seus jornalistas e na perceção dos leitores. As variações identificadas permitem, num segundo momento, refletir sobre o poder que os dados têm – ou podem ter – no seio digital e sobre o seu efeito regenerador da identidade e credibilidade jornalística. Para os jornalistas do Expresso, os dados promovem a descoberta de informação nova e potencializam a vertente multimédia do digital. E para os leitores, os dados podem melhorar a exatidão das histórias e, sobretudo, reanimar a sua confiança no jornalismo.

**Palavras-chave:** digital, Internet, tecnologias, sistemas de informação, jornalismo de dados, bases de dados, jornal Expresso





## **ABSTRACT**

If journalism has a power is the power to adapt to circumstances, to adversities, to technological advances, to new ways of transmitting information. From its adaptation to the Internet was born digital journalism, which, on the one hand, brings many possibilities, but on the other, creates an environment of instability that affects the credibility's journalism. And from the need to understand the billions of data, that are generated every day in the digital world, flourishes data journalism. This practice, which extracts stories from a lot of data, has become a new international journalistic trend. There are several media integrate data journalism in their newsrooms. And the newspaper Espresso is no exception. In order to understand this phenomenon, the six months of internship in the Espresso Digital, the readings, the interviews and the questionnaires – gathered in this report – look for to explore the changes that the integration of data journalism generates in this newsroom, in the productive routines of its journalists and on the readers' perception. The variations identified allow, in a second moment, to reflect about the power the data have – or may have – in the digital world and about its regenerative effect on journalistic identity and credibility. For Espresso journalists, the data promote the discovery of new information and enhance multimedia stories. And for readers, the data can improve the accuracy of the stories and, above all, revive their confidence in journalism.

**Keywords:** digital, Internet, technologies, information systems, data journalism, databases, newspaper Espresso



## ÍNDICE GERAL

INTRODUÇÃO .....	1
1 O Jornalismo e as novas tecnologias .....	5
1.1 Como as tecnologias e a Internet entraram no Jornalismo.....	5
1.2 O impacto das tecnologias no seio jornalístico .....	7
1.3 Um caminho lado a lado: como o Jornalismo abraçou os Dados .....	14
1.3.1 O que o Jornalismo de Dados traz de novo .....	18
1.3.2 O Jornalismo de Dados enquanto fluxo de trabalho .....	21
1.3.2.1 Da análise de dados nascem as histórias .....	23
1.3.2.2 Aplicações de notícias: uma nova forma de contar histórias .....	27
2 A integração do Jornalismo de Dados e as mudanças subjacentes.....	29
2.1 As mudanças nas redações .....	29
2.2 Um novo perfil emergente: o jornalista de dados .....	32
2.3 Uma viagem interativa pelos dados: o que muda para o público.....	35
3 O semanário que se reinventa: o jornal Expresso .....	37
3.1 Traços da história do jornal Expresso .....	38
3.2 Do Papel ao Digital: uma breve caracterização .....	40
3.2.1 Estratégias de negócio do Expresso Digital.....	42
3.3 As rotinas produtivas: da seleção de temas à publicação.....	43
3.3.1 Como é produzida a rubrica 2:59? .....	46
4 Seis meses no seio do jornal Expresso.....	49
4.1 Objetivos do estágio.....	49
4.2 Diário de Bordo.....	51
4.2.1 Acolhimento.....	51
4.2.2 Projetos Desenvolvidos.....	52
4.2.3 Desafios e Aprendizagens.....	63

4.2.4	Principais tópicos de investigação .....	65
4.2.5	Conclusão do Estágio.....	68
5	Investir e inovar em jornalismo de dados: o caso do jornal Expresso .....	69
5.1	Como é que os jornalistas do Expresso abraçam o jornalismo de dados? .....	70
5.1.1	Recolha, seleção e métodos de análise .....	70
5.1.2	Caracterização sociodemográfica e profissional das amostras .....	72
5.1.3	O jornalismo de dados aos olhos de duas gerações distintas .....	73
5.1.4	Discussão de resultados .....	83
5.2	Qual é a perceção dos leitores do Expresso em relação às histórias baseadas em dados? .....	85
5.2.1	Recolha de dados, metodologia e objetivos.....	85
5.2.2	Dados sociodemográficos .....	85
5.2.3	O jornalismo de dados aos olhos dos leitores dos jornais digitais generalistas.....	86
5.2.4	Um caminho possível: reforçar a aposta em jornalismo de dados.....	87
	CONCLUSÃO .....	89
	REFERÊNCIAS BIBLIOGRÁFICAS .....	93
	APÊNDICES .....	99
	APÊNDICE 1. Entrevista ao editor adjunto do Expresso Diário José Cardoso .....	100
	APÊNDICE 2. Modelação do processo de produção de notícias no jornal Expresso: desde a ideia inicial até à publicação .....	109
	APÊNDICE 3. Modelação do processo de produção da rubrica 2:59 .....	110
	APÊNDICE 4. Visualizações referentes às perceções dos leitores quanto ao jornalismo de dados .....	111
	APÊNDICE 5. <i>Short Paper</i> aceite na 13ª edição da Conferência Ibérica de Tecnologias e Sistemas de Informação (CISTI).....	114

## ÍNDICE DE FIGURAS

<b>Figura 1.1</b> - A Pirâmide Invertida do Jornalismo de Dados .....	21
<b>Figura 4.1</b> – Tabela e gráfico dinâmicos que mostram a evolução do número mulheres e homens eleitos presidentes desde 1976 até 2017 .....	54
<b>Figura 4.2</b> – Número total mandatos (femininos e masculinos) por partidos e coligações de autarcas (diferenciado por género).....	54
<b>Figura 4.3</b> – Número de mandatos femininos por concelho (à esquerda) e evolução do número de mandatos conduzidos por mulheres desde 1976 até 2017 (à direita) .....	55
<b>Figura 4.4</b> – Interface do <i>site</i> do Expresso da história sobre as mulheres autarcas (a partir da qual é possível aceder à história integral publicada no Expresso Diário).....	57
<b>Figura 4.5</b> – Interface do <i>site</i> do 2:59 para explicar o mercado do café português .....	63

## ÍNDICE DE TABELAS

<b>Tabela 4.1</b> – Correspondência entre objetivos e capítulos de desenvolvimento.....	50
<b>Tabela 4.2</b> – Sistemas de informação utilizados no processo de jornalismo de dados ..	66
<b>Tabela 5.1</b> – Percepção do grau de dificuldade dos jornalistas mais e menos experientes do jornal Expresso relativamente às tarefas incluídas no processo de jornalismo de dados (em % face ao total de respostas de cada grupo) .....	74
<b>Tabela 5.2</b> – As capacidades que os jornalistas de dados devem desenvolver na perspetiva dos jornalistas do Expresso que iniciaram a carreira antes e depois de 2000.....	79

## ÍNDICE DE GRÁFICOS

<b>Gráfico 5.1</b> – Frequência de produção de notícias baseadas em dados pelos jornalistas que iniciaram a carreira antes e depois do ano 2000 .....	73
<b>Gráfico 5.2</b> – Principais fontes de recolha de bases de dados ou relatórios técnicos utilizadas pelos jornalistas do Expresso que iniciaram a carreira antes e depois do ano 2000 (em % face ao total de respostas de cada grupo) .....	76
<b>Gráfico 5.3</b> – Sistemas de informação que suportam o processo de jornalismo de dados utilizados pelos jornalistas do Expresso que iniciaram a carreira antes e depois de 2000 .....	77
<b>Gráfico 5.4</b> – Perceção dos jornalistas do Expresso, que iniciaram a carreira antes e depois do ano 2000, quanto ao grau de integração do jornalismo de dados na redação. 80	
<b>Gráfico 5.5</b> - Aspetos que a integração do jornalismo de dados pode melhorar segundo a perspetiva dos jornalistas que iniciaram a carreira antes e depois de 2000 .....	81



## **Lista de abreviaturas, acrónimos e siglas**

ARPANET - *Advanced Research Projects Agency Network*

BBC - *British Broadcasting Corporation*

CISTI – Conferência Ibérica de Tecnologias e Sistemas de Informação

ENIAC - *Electronic Numerical Integrator and Computer*

GPS - *Global Positioning System*

HTML – *Hypertext Markup Language*

SIC - Sociedade Independente de Comunicação

SQL – *Structured Query Language*

TVI - Televisão Independente

VOIP - *Voice Over Internet Protocol*

WWW – *World Wide Web*

## **INTRODUÇÃO**

### **Enquadramento**

São inegáveis as transformações que as novas tecnologias trazem para o mundo contemporâneo. No jornalismo, em particular, são enormes. Hoje, todos – ou quase todos – os processos das redações são suportados por tecnologias e sistemas de informação.

A aposta no jornalismo digital – que emergiu no seio da Internet – é cada vez mais robusta. São várias as possibilidades que o meio digital gera: passa a ser possível construir histórias multimédia, interligá-las e ainda contactar os leitores e as fontes por novas vias. Mas também são vários os desafios que o jornalismo enfrenta. O aumento da concorrência no seio digital origina uma instabilidade económica tal, que submete os valores jornalísticos a uma lógica comercial, abalando por completo a identidade dos jornalistas e, por conseguinte, a sua credibilidade aos olhos do público. Perante este cenário, os *media* procuram uma solução capaz de devolver a integridade ao jornalismo, recuperar a confiança dos leitores e criar estabilidade financeira nas redações. São várias as teorias, as fórmulas estudadas. Mas, até então, não se avistam mudanças significativas no horizonte.

Poderá o jornalismo de dados ser uma das hipóteses testadas? Este processo – que procura encontrar novas histórias na imensidão de dados todos os dias gerados, com o auxílio de sistemas e tecnologias de informação – já é considerado uma tendência jornalística internacional. São cada vez mais os *media* que o adotam. E já há vários projetos digitais de sucesso, como a *DataStore* do *ProPublica* e a rubrica 2:59 do Expresso Digital. Mas o que traz de novo esta prática jornalística que explique tamanha adesão? Que mudanças surgem aquando a sua integração? O que muda para os jornalistas? E para os leitores?

Estas questões refletem a necessidade de conhecer de perto o impacto que os dados têm – ou podem ter – no mundo do jornalismo. E, para tal, integrar uma redação que hoje aposta no jornalismo de dados revela-se o melhor caminho, o caminho que levou a investigadora a estagiar durante seis meses no seio do Expresso Digital.

### **Objetivos**

Compreender o poder que o jornalismo de dados tem – ou pode ter – no mundo digital foi a principal força motriz que despoletou o desenvolvimento do presente relatório. Com o intuito de reunir a informação necessária para atingir este objetivo, a investigadora delineou uma série de metas. Estudar como o jornalismo se molda às novas tecnologias,

bem como os impactos associados. Investigar as origens do jornalismo que dos dados traz novas histórias, a sua ascensão e as várias fases do processo. Explorar as mudanças que esta prática pode despoletar nas redações, nas rotinas dos jornalistas e na vida dos leitores. Todas estas metas permitem enquadrar teoricamente as temáticas em estudo, mas a necessidade de conhecer de perto como tudo se concretiza levou a investigadora a integrar uma redação que investe no jornalismo de dados: o Expresso Digital. Os seis meses de estágio neste jornal permitiram alcançar outras metas definidas. Compreender como uma redação integra projetos digitais em que o jornalismo de dados é protagonista – no caso do Expresso, a rubrica 2:59, bem como outros conteúdos baseados em dados. Estudar as rotinas dos jornalistas do Expresso que adotam esta prática, bem como a perceção dos seus leitores, a fim de compreender o que nelas mudam. O estudo do caso particular do Expresso Digital procura, assim, ser um ponto de partida para compreender um fenómeno maior: a crescente integração do jornalismo de dados nos *media* internacionais.

## **Metodologia**

Para reunir todas as informações necessárias para alcançar os objetivos supracitados foram utilizadas diferentes técnicas de recolha de dados. O recurso à literatura, a observação participante, a realização de entrevistas e de questionários foram as aplicadas.

Para compreender o contexto em que o jornalismo se adaptou às novas tecnologias e o ambiente em que floresce o jornalismo de dados, o recurso à literatura científica revelou-se fundamental. Mas para perceber de perto como uma redação hoje trabalha com as novas tecnologias e lida com os dados, as entrevistas aos editores do Expresso e a observação participante, apenas possível em contexto de estágio, revelaram-se fundamentais. Esta última técnica, aplicada durante seis meses no seio do jornal Expresso, permitiu à investigadora não só observar e analisar as rotinas da redação, como também fazer parte delas. E permitiu ainda pôr em prática o processo de jornalismo de dados, não só integrando a rubrica 2:59, mas também construindo histórias baseadas em dados.

Para estudar as mudanças subjacentes ao fenómeno internacional que é a integração massiva do jornalismo de dados nas redações pelo mundo, o estudo do caso ilustrativo do Expresso Digital foi a estratégia de pesquisa utilizada, a qual foi baseada não só na observação participante, como também em questionários que foram direccionados aos jornalistas do jornal, bem como aos leitores.

## **Estrutura do Trabalho**

O presente relatório está estruturado em cinco capítulos distintos, mas complementares. O primeiro tem por base a contextualização das temáticas em estudo. O segundo explora as mudanças organizacionais que estas geram. O terceiro aprofunda a organização em estudo: o jornal Expresso. E o quarto procura condensar as principais aprendizagens vividas em contexto de estágio neste jornal. O quinto e último capítulo é referente ao estudo do caso do jornal Expresso, enquanto *media* que investe em jornalismo de dados.

O primeiro capítulo resulta da investigação teórica e procura contextualizar, em três subcapítulos organizados segundo uma lógica temporal, as temáticas em estudo. O primeiro subcapítulo regressa ao momento em que as novas tecnologias e a Internet entraram no mundo do jornalismo. E o segundo, por sua vez, explora as várias possibilidades e armadilhas geradas por este mar digital em que o jornalismo mergulhou. O terceiro subcapítulo viaja pelas origens do jornalismo que dos dados traz novas histórias, passando pelos momentos que despoletaram a sua evolução até às várias fases que o compõem o processo. No capítulo segundo estão reunidas as mudanças geradas pela integração do jornalismo de dados na estrutura das organizações mediáticas, nas rotinas dos jornalistas e na perceção dos leitores – correspondendo cada ator a um subcapítulo particular. O terceiro capítulo procura conhecer de perto o jornal que acolheu a investigadora durante seis meses em três níveis distintos – sendo que cada um corresponde a um subcapítulo. Ao nível dos momentos que marcam a história e a evolução do jornal Expresso. Ao nível das plataformas particulares que hoje fazem respirar o jornal e das estratégias de negócio aplicadas no seio digital. E também ao nível dos processos produtivos e, em especial, da produção da rubrica 2:59. O quarto capítulo reúne os objetivos inicialmente propostos, num subcapítulo, e o diário de bordo, noutro, no qual estão reunidas as experiências, as aprendizagens e as dificuldades vividas no seio do jornal Expresso durante seis meses de estágio. O quinto e último capítulo diz respeito ao estudo de caso, que analisa as mudanças que a integração do jornalismo de dados traz para os jornalistas do Expresso, num subcapítulo, e para os seus leitores noutro.

## **Contributos**

As perceções dos leitores quanto às histórias baseadas em dados, recolhidas através de questionários no âmbito deste relatório, deram origem a um outro trabalho, a um *short paper* intitulado “Um Olhar sobre o Jornalismo de Dados” (apêndice 5) que foi aceite na 13ª Conferência Ibérica de Tecnologias e Sistemas de Informação realizada em 2018.



## **1 O Jornalismo e as novas tecnologias**

### **1.1 Como as tecnologias e a Internet entraram no Jornalismo**

Tudo começou em 1946, em meados do século XX, quando os dois cientistas norte-americanos John Eckert e John Mauchly anunciaram a primeira máquina digital capaz de armazenar e processar informação, o *Electronic Numerical Integrator and Computer* (ENIAC) (Fernandes, 2000). Esta foi a invenção que transformou o mundo: a sociedade, a economia, a política, a saúde, a educação. Nenhuma área ficou indiferente. E o jornalismo não foi exceção.

Os anos 70 do século passado marcam uma série de avanços e inovações tecnológicas. A capacidade de armazenar informação nos computadores foi renovada pela *International Business Machines* (IBM) e os processadores da *Intel Corporation* tornaram-se mais rápidos. Os computadores pessoais também chegaram ao mercado: Ed Roberts e Bill Yates criaram o *Altair 8800* e Steve Jobs e Steve Wozniak o *Apple I*. Os computadores, já aprimorados, não demoraram a entrar no mundo jornalístico. A produção de notícias pela via mecânica – isto é, com auxílio da máquina de escrever – viu os seus dias contados. As potencialidades dos computadores foram rapidamente percebidas pelos jornalistas, pelos editores e pelos gestores dos meios de comunicação norte-americanos, o que se traduziu num investimento tecnológico (Moherdaui, 2017).

Também nos Estados Unidos da América (EUA) surge a *Advanced Research Projects Agency Network* (ARPANET), a primeira rede que interligou os militares americanos em plena Guerra Fria, em 1969. Esta rede, criada pelo Governo dos EUA, visava proteger os militares de um possível ataque da União Soviética. E foi esta pequena rede que deu origem à que hoje interliga o mundo, a Internet. Os meios de comunicação não tardaram a perceber o seu potencial e, também na década de 70 do século XX, começaram a fazer os primeiros testes de difusão de informação por este novo meio. A *British Broadcasting Corporation* (BBC) deixou-se inspirar pela Internet e começou a transmitir novos formatos multimédia, nomeadamente o videotexto. O jornal norte-americano *The New York Times* criou, por sua vez, criou o *InfoBank* onde eram publicados artigos do jornal por via do sistema *Biennial Reporting System*. Mais tarde, em meados de 1990, surge a ferramenta que atualmente mais se utiliza para aceder à Internet, a *World Wide Web* (WWW). Com a invenção do engenheiro britânico Tim Berners-Lee passou a ser possível conectar o mundo através deste novo meio (Moherdaui, 2017).

Este panorama despoletou toda uma fase de mudança nas redações pelo mundo, que começaram a adotar novas tecnologias computacionais e a apostar fortemente neste novo meio de transmissão de informação. Tantos os jornais como as rádios e as televisões criaram as suas páginas na *web*, onde disponibilizavam gratuitamente os mesmos conteúdos já publicados nas respetivas plataformas tradicionais (Canavilhas, 2005). Em Portugal, os meios de comunicação iniciaram este processo entre 1995 e 1998. O Jornal de Notícias e o jornal Público foram os primeiros jornais diários generalistas a investir numa edição digital, em 1995. No mesmo ano, também a Rádio e Televisão de Portugal (RTP), inaugurou a sua página *web*. Em 1996 foi a vez dos outros dois canais televisivos de sinal aberto, da TVI (Televisão Independente) e da SIC (Sociedade Independente de Comunicação), e também da rádio TSF (Bastos, 2007). Em meados de 1997, o jornal Expresso anunciou, na sua edição em papel, que passaria a publicar as suas edições integrais na sua nova página *web*. E este, segundo Bastos (2007), foi “*o primeiro semanário português a estar presente na Internet*”.

Abraçar aquela que prometia ser a rede que conectaria o mundo foi um importante passo que marcou o princípio da revolução digital da informação. Ainda assim, nos primeiros tempos, a maioria das redações portuguesas não evidenciavam renovações expressivas. Nas páginas *web* eram divulgados os conteúdos que já haviam sido publicados nas plataformas tradicionais dos respetivos *media*. Em exclusivo para a Internet, apenas eram publicadas notícias de última hora e atualizações de temas importantes. O meio de transmissão da informação era diferente, mas o modo de fazer jornalismo era o mesmo. Os modos de narrar eram os mesmos. E os processos de produção noticiosa pouco mudavam. Já as potencialidades e os perigos da *web* estavam longe de ser percebidos (Canavilhas, 2005).

## **1.2 O impacto das tecnologias no seio jornalístico**

O impacto das novas tecnologias e da Internet no seio jornalístico começa a fazer-se sentir quando as suas valências começam a ser exploradas, quando esta nova vertente digital do jornalismo começa a ganhar força nas redações. A partir da viragem do milénio, este impacto começa a traduzir-se em mudanças não só nas estruturas das organizações jornalísticas, como também nas rotinas produtivas dos jornalistas. A força da Internet renova também a relação entre jornalistas e o público, que passa a ter dois sentidos, e transforma, por completo, a forma comunicar dos cidadãos na sociedade contemporânea.

Foi na viragem do milénio que se assistiu a um forte investimento na *web* por parte dos meios de comunicação social em Portugal. Alguns, inclusive, investiram numa versão somente digital, como foi o caso do Diário Digital (Bastos, 2007) – que fechou portas quase 18 anos depois, em janeiro de 2017. Ao contrário dos outros *media*, este jornal generalista apostou na produção de conteúdos exclusivamente para o meio digital, tirando partido de algumas das suas características. As notícias passaram a disponibilizar ligações a outros conteúdos – o denominado hipertexto. As histórias, para além de serem contadas através de textos e imagens, passaram a incorporar áudios e vídeos. Passaram, portanto, a ter traços multimédia. O Diário Digital estimulou também a comunicação entre leitores e jornalistas através da *web*, promovendo, assim, uma comunicação em dois sentidos e não apenas um. A essência da interatividade no meio digital começa, então, a ser percebida (Canavilhas, 2005).

Estas potencialidades da rede global foram sendo, passo a passo, reconhecidas pelos meios de comunicação, gerando uma onda de mudança. Novas técnicas de redação foram desenhadas, bem como processos de produção noticiosos. E novas formas de contactar as fontes e de recolher informação surgiram. Canavilhas (2003) recorda que este processo de descoberta e adaptação do jornalismo a um novo meio de difusão de informação – neste caso, a Internet –, não é novo na sua história. Em tempos, o jornalismo adaptou-se à rádio e à televisão, o que introduziu narrativas jornalísticas particulares. Na viragem do milénio, foi a vez de o jornalismo se adaptar ao mundo digital, um meio onde tudo muda.

Para os jornalistas, a Internet desencadeou – e ainda desencadeia – um conjunto de mudanças nas suas rotinas produtivas. No que diz respeito à produção de notícias, os jornalistas passam a ter o poder de incorporar na mesma história vários meios – como texto, vídeo, gráficos, infografias (interativas ou não), áudios, ilustrações, fotografias,



entre outros –, algo que antes não era possível. As limitações de espaço também deixam de ser uma preocupação, pois na Internet o espaço pode ser, se quisermos, infinito e interligado (Canavilhas, 2006). Ao potenciar estas características do meio digital, os jornalistas libertam-se da tradicional pirâmide invertida – que prioriza a informação mais importante e que responde às seis questões que compõem o *lead* (o quê, quem, quando, onde, como e porquê) – e passam a abraçar outras superestruturas, como é o caso da pirâmide deitada. Esta superestrutura, em particular, organiza a informação por quantidade e não por importância. Numa primeira fase, os jornalistas redigem um resumo do sucedido – o correspondente ao *lead* –, e à medida que contam a história introduzem outros conteúdos – através de hiperligações, por exemplo –, aumentando, deste modo, o número de informações que explicam e contextualizam o acontecimento, as quais podem ser comunicadas através de vários meios (Canavilhas, 2006). Deste modo, os jornalistas deixam de definir um percurso de leitura e passam a dar a oportunidade aos leitores de criarem a seu, à sua medida. Cada leitura passa, então, a ser subjetiva e guiada de acordo com os interesses pessoais de cada um. A natureza da Internet permite também que os leitores possam interagir não só com a história, mas também com os jornalistas e com outros leitores, através de comentários, por exemplo (Canavilhas, 2003).

No que diz respeito à recolha de informação, os jornalistas viram o leque de opções a aumentar. Para além da observação direta dos acontecimentos – uma tradicional forma de obter informação –, os jornalistas passam a ter a possibilidade de pesquisar rapidamente informações na Internet, nas páginas oficiais de organizações governamentais e não governamentais, de associações e de outros *media*. O número de alternativas para realizar entrevistas a especialistas também aumenta, pois passa a ser possível contactá-los via *email* e através das redes sociais, por exemplo, para além de pessoalmente e via telefónica. As possibilidades geradas por este novo meio são tantas, que passa a ser possível construir uma notícia de raiz e publicá-la sem sair da redação. Mas será que estas novas formas de recolher informação se tornaram as principais escolhas dos jornalistas? Até que ponto as práticas tradicionais de recolha de informação são substituíveis? Num contexto organizacional, onde as pressões de tempo vigoram, pesquisar informações e entrevistar especialistas rapidamente é uma mais valia para os jornalistas (Gomes, 2008). Mas há aspetos que se perdem, há informações que não são possíveis de recolher através da Internet. O ambiente que envolve um acontecimento somente é percebido quando vivenciado, quando um jornalista se desloca ao local e o observa diretamente, quando o

jornalista é uma testemunha privilegiada de um evento. E toda a comunicação não verbal – como os gestos e a postura – também se perde quando o jornalista deixa de entrevistar as suas fontes presencialmente. Além disso, nem sempre é fácil encontrar informações credíveis mergulhando na Internet, que reúne bilhões de dados e informações gerados pelos quatro cantos do mundo (Gomes, 2008).

Os novos espaços sociais, que emergiram no mundo digital, vieram mudar para sempre as rotinas dos jornalistas como também as dos cidadãos em geral. As redes sociais – como o *Facebook*, *Twitter* e o *Instagram* – são hoje um importante trunfo para o *media*, por várias razões, segundo Pavlik John (2014). É no seio destas redes que é possível identificar e entrevistar especialistas. E é nelas que é possível identificar novos temas de discussão pública. As redes – onde todos os dias são publicados milhares de conteúdos – são hoje uma fonte rica em informação em primeira mão. Há registos de acontecimentos que são todos os dias publicados pelo público nestas redes – quer em formato de vídeo, áudio e imagem, quer em forma de testemunhos escritos –, os quais podem ilustrar ou complementar uma notícia. Quando o público assume este papel, o de registar e divulgar pedaços de uma verdade de interesse público, erguer-se o denominado jornalismo de cidadão (Pavlik John, 2014). Para o público, estas redes revelam-se autênticos espaços de discussão pública. As redes sociais vieram reforçar o papel cidadãos enquanto atores sociais individuais, com ideais e opiniões próprias (Pissarra Esteves, 2007). Vieram dar um novo fôlego à voz do público, que durante anos esteve submissa à comunicação de massas. Se antes a comunicação se traduzia na relação de “*um para muitos*” – sendo que um *media* comunicava para muitas pessoas –, com chegada a Internet esta relação muda e passa a ser “*de muitos para muitos*” (Gomes, 2008). Isto quer dizer que o público ganha voz e ganha também a capacidade de produzir e publicar conteúdos nesta nova rede global. É através do meio digital, e em especial no seio das redes sociais, que o público – cada vez mais informado e exigente – dá o seu parecer sobre os conteúdos jornalísticos produzidos. “*O feedback da audiência às notícias publicadas passou a ser uma realidade*” neste novo meio, refere Gomes (2008). Mas será que os jornalistas prestam atenção às inquietações do público e às suas reações? Será que comunicam com os seus leitores? De que forma os dados gerados pelo público podem ser usados em prol de um jornalismo de qualidade? Estas são algumas questões que se podem colocar para – quem sabe – melhor compreender a importância da interatividade, entre jornalistas e o público, no seio jornalístico.

Neste contexto digital, as tecnologias tornam-se parte integrante das redações e passam a fazer parte das rotinas dos seus profissionais. O jornalismo tem o potencial de se reinventar a cada nova notícia. Os jornalistas libertam-se das limitações de espaço e adotam novas técnicas de redação, procurando potencializar as características próprias deste novo meio de difusão, a Internet. E os leitores, por sua vez, experimentam uma autêntica viagem multimédia e veem a distância, até então, sentida entre si e os meios de comunicação encurtada. Todos estes aspetos positivos sumariados são, sem dúvida, importantes para o desenvolvimento do jornalismo digital enquanto prática contemporânea. O panorama parece animador. Aos olhos dos jornalistas “*a Internet é vista como um novo mundo, essencialmente cor-de-rosa*”, refere Gomes (2008). Os benefícios da *web* são sublinhados e, por outro lado, os perigos da mesma parecem desvanecer na perspetiva destes profissionais. “*A correria rumo ao virtual causou uma espécie de cegueira em tons de rosa*”, sublinha Gomes (2008). Mas tal como as rosas, a Internet – enquanto ferramenta de trabalho e novo meio para difusão de informação – também tem espinhos, que devem ser identificados. Olhar para estes espinhos é ter em conta um outro lado, um lado que mostra como a sustentabilidade dos *media* está em perigo. Que mostra como os interesses comerciais se sobrepõe aos valores jornalísticos. Um lado que mostra como os jornalistas estão subordinados a esses mesmos interesses e mudam, por isso, a forma como relatam a verdade e como a procuram.

Um fenómeno que desencadeou uma onda de mudança na estrutura das organizações noticiosas diz respeito à revolução no acesso à informação, gerada pela Internet. O número de fontes informativas – mediáticas ou não – aumentou exponencialmente no meio digital. E, por conseguinte, hoje há mais informação disponível ao público do que alguma vez houve. Este panorama extinguiu para sempre o monopólio de produção de conteúdos informativos, até então, detido pelos *media* e criou uma nova estrutura de mercado, onde os *media* concorrem com outras organizações não noticiosas (Picard, 2013) e onde o público, por sua vez, tem o poder de produzir conteúdos (Corrêa & Lima, 2013). Face a esta dispersão do mercado, os *media* viram a sua estabilidade económica abalada. Este cenário resultou em sérias quedas nas vendas de jornais, nas audiências das televisões e no número de ouvintes das rádios (Garcia, 2009). O semanário Expresso, por exemplo, apesar de ser o jornal mais vendido no país no terceiro bimestre de 2018 com mais de 86 mil exemplares vendidos, sentiu fortes quedas nas vendas nas últimas décadas. A venda de exemplares caiu 22,5% entre o terceiro bimestre de 2010 e o período

homólogo de 2018. E 21,6% entre os terceiros bimestres de 2000 e de 2010. Isto quer dizer que desde início do milénio até à atualidade, o Expresso sentiu uma queda nas vendas na ordem dos 40%, o equivalente a mais de 55 mil exemplares (APCT, 2018). Como consequência das sucessivas quedas, os *media* veem também as receitas publicitárias a diminuir dia após dia. O declínio de leitores, telespectadores e de ouvintes significa, na ótica das empresas, um menor alcance dos seus conteúdos publicitários, o que as leva a procurar outras alternativas para promover os seus produtos e serviços, nomeadamente no seio digital. Como se não bastasse, as plataformas digitais dos *media* não são – ainda – financeiramente sustentáveis (Coelho, 2013). Apesar dos vários modelos de receita testados ao longo do milénio – como é o caso da publicidade em diferentes formatos, do *freemium*, do *pay wall* e dos fundos de investimento –, a fórmula mágica que tornará as páginas *web* dos *media* sustentáveis em termos financeiros ainda está para ser descoberta. Ainda assim, estas experiências já revelaram, por exemplo, que para maximizar as receitas publicitárias na *web*, o acesso aos conteúdos – pelo menos em parte – deve ser gratuito, já que estas receitas dependem diretamente do número de leituras obtidas (Corrêa & Lima, 2013). Tendo em conta esta lógica, os gestores das organizações mediáticas começam a adotar estratégias a nível editorial que passam por sobrevalorizar o número de leituras dos conteúdos em detrimento da qualidade e do rigor jornalístico. Passam a encarar as notícias como produtos, que devem, numa primeira instância, atrair o público a que se destinam e, em segundo plano, informar e criar conhecimento. A essência do jornalismo vê-se abalada, diminuída, e surge a verdadeira crise que o jornalismo enfrenta, a de valores (Garcia, 2009).

A reestruturação das organizações mediáticas – desencadeada pela crise financeira – refletiu-se diretamente nas redações e nas rotinas dos seus profissionais. A convergência de meios – humanos e técnicos – fez-se sentir no seio dos grupos mediáticos. E as vagas de despedimento tornaram-se cada vez mais frequentes. Os jornalistas que resistem veem não só as suas condições de trabalho a degradar-se, como também a agenda mediática subordinada a interesses estritamente económicos. Assiste-se, portanto, a uma inversão de relevância de temas ditada pelos gestores dos *media*, que Garcia (2009) descreve do seguinte modo:

*“A exploração da faceta agradável, divertida e evasiva dos jornais, que tem porventura raízes na arte de contar histórias de ‘interesse humano’ (dramas, intrigas, mexericos e rumores), que – não esqueçamos – são também componente*

*da vida social e do jornalismo, encaminha-se para hegemonizar este sector”*  
(Garcia, 2009:31).

Segundo esta dinâmica comercial, os bons conteúdos jornalísticos são aqueles que atraem os leitores, que promovem vendas e visualizações no seio digital. São os conteúdos prazíveis, curtos e de fácil leitura, passíveis de serem publicados rapidamente (Correia, 2006). E não aqueles que revelam verdades desconhecidas, aqueles que relatam uma realidade que influencia – ou pode influenciar – diretamente a vida de cada um.

*“Os jornalistas estão enquadrados em normas deontológicas que os impedem de colocar as vendas ou a merca diversão acima do interesse/serviço público”*, salienta Graça (2007). Mas, sendo assalariados, os jornalistas ficam subordinados às escolhas editoriais, que procuram rentabilizar a informação e, que, por isso, acabam por espezinhar a identidade jornalística. Isto quer dizer que as lógicas de mercado se sobrepõem à ética e aos valores orientadores da ação jornalística. Os jornalistas veem, assim, a sua liberdade na escolha editorial reduzida, assim como a sua autonomia e independência no exercício da profissão (Coelho, 2013). Neste contexto, os jornalistas não deixam de procurar a verdade – um valor central no jornalismo –, mas fazem-no com constrangimentos e limitações. Coelho (2013) salienta que *“no meio digital, não há receitas que suportem as deslocações dos jornalistas para realizar reportagens em campo e ter contacto com o real”*, procedimentos que fazem parte das raízes da profissão. Em resultado, o jornalismo perde profundidade e contexto. E assiste-se à sedentarização da profissão (Coelho, 2013). Perante o imediatismo, ainda mais presente no seio digital, a verificação de factos, a avaliação de fontes de informação, a apresentação dos dois lados da questão e a exatidão e o rigor da informação – procedimentos que constroem a credibilidade jornalística (Traquina, 2007) – são, por vezes, esquecidos. É neste contexto que o papel do jornalismo enquanto guardião da verdade e defensor da opinião pública é questionado (Coelho, 2013). E esta é uma ponta do iceberg da crise de identidade profissional que hoje paira sobre o jornalismo e sobre os seus profissionais (Garcia, 2009).

Uma outra ponta deste iceberg diz respeito à relação entre o público e os *media*, que tem vindo a deteriorar-se dia após dia. São várias as razões que explicam este divórcio. Os leitores reconhecem que a qualidade das notícias é cada vez menor. Esta perceção do público pode, segundo Coelho (2013), estar intrinsecamente relacionada com o despedimento de jornalistas veteranos que não só conquistaram relações próximas com fontes importantes como também possuem um conjunto de saberes profundos no campo

jornalístico – como é o caso do saber reconhecer notícias (tendo por base critérios de noticiabilidade e valores-notícia), saber identificar e seleccionar factos e saber narrar histórias de forma compreensível para o público (Traquina, 2007). Além disso, os leitores percebem que hoje as notícias são, mais vírgula menos vírgula, todas iguais e reconhecem também a lógica comercial por detrás da sua publicação (Correia, 2006). E identificam igualmente uma série de incongruências nas notícias apresentadas, que derivam da falta de rigor no trabalho jornalístico. As notícias de última hora, publicadas sob curtos espaços de tempo, são suscetíveis de tais erros. A confirmação e verificação de factos ficam, por vezes, esquecidas em detrimento do imediato. Mas a falta de ação neste sentido poderá levar, por exemplo, a que uma notícia falsa seja difundida por vários meios de comunicação que, até então, eram considerados credíveis aos olhos do público. Todo este panorama gera um ambiente de desconfiança dos leitores nos conteúdos jornalísticos e no próprio jornalismo (Souza, 2017).

O jornalismo vive um momento de pura transformação desencadeada pelo mundo digital. E esta transformação dá-se em dois sentidos. Um deles abre caminho a um mundo de possibilidades, que têm o potencial para desenvolver e melhorar o jornalismo. E outro afeta gravemente a estrutura financeira dos *media*, gerando um clima de instabilidade. As decisões editoriais tomadas pelos gestores das organizações mediáticas para fazer face a esta crise deram lugar a outras. A uma crise de valores jornalísticos, que por sua vez originou uma crise de confiança no jornalismo. E, até então, este clima não dá sinais de mudança.

### **1.3 Um caminho lado a lado: como o Jornalismo abraçou os Dados**

Os dados sempre fizeram parte do jornalismo. Mas não sempre da mesma forma, nem com a mesma intensidade. O modo de recolher dados mudou. As metodologias de limpeza e análise de dados também não são as mesmas. E as tecnologias que estão à disposição dos jornalistas são completamente diferentes.

No século XIX, os jornalistas recolhiam, eles próprios, os dados. Para conhecer o número de alunos que realmente recebiam educação gratuita na cidade de *Manchester*, os jornalistas do diário britânico *The Guardian* investigaram não só quantos alunos estudavam em cada escola, como também as despesas associadas. Esta reportagem, publicada em 1821, é um dos primeiros trabalhos jornalísticos baseados em dados de que há registo (Bounegru, 2012). Nesta época, os jornalistas passavam semanas, meses e até anos a recolher e analisar manualmente um enorme conjunto de dados. A probabilidade de erro era considerável e, conseqüentemente, o rigor jornalístico questionável. Apesar das desvantagens, a necessidade de encontrar novas abordagens, de investigar temas e de descobrir informações pertinentes levou os jornalistas a não desistir e a continuar a investir nesta vertente jornalística, que centra os dados (Teixeira & Junior, 2012).

O aparecimento das novas tecnologias digitais, em meados dos anos 50 do século XX, foi uma verdadeira lufada de ar fresco para esta prática, para o jornalismo que abraça os dados. A primeira notícia baseada em dados elaborada com recurso a um computador foi publicada em 1952, pela cadeia de rádio e televisão norte-americana CBS, e procurava prever o resultado das eleições presidenciais dos Estados Unidos da América (EUA) naquele ano. Analisar grandes quantidades de dados com recurso a um computador para descobrir e melhorar histórias tornou-se uma prática, com uma metodologia própria, a qual foi batizada de Reportagem Assistida por Computador. Durante anos a fio, este fluxo de trabalho foi, sobretudo, empregado nos EUA por jornalistas de investigação, que procuravam defender o interesse público, supervisionando o poder de forma independente (Bounegru, 2012). As tecnologias e sistemas de informação vieram, assim, apoiar a recolha e análise de grandes quantidades de dados e trouxeram benefícios claros: o tempo de execução das tarefas foi encurtado e os resultados gerados tornaram-se mais rigorosos.

Esta abordagem jornalística ganha um novo alento na década de 70. O jornalista Philip Meyer – com o seu livro intitulado *The New Precision Journalism* – propõe aliar os métodos das ciências sociais às práticas jornalísticas, para encontrar verdades nos dados

e, por conseguinte, produzir um jornalismo mais próximo da objetividade (Bounegru, 2012). A introdução destes novos métodos nas redações veio desafiar os jornalistas a dominar saberes que vão para lá da arte de bem escrever. Aprender a aplicar métodos estatísticos das ciências sociais – como definir amostras e realizar testes de hipóteses – para recolher, trabalhar e compreender os dados, passou a ser crucial. E saber lidar com tecnologias e sistemas de informação, que auxiliam o processo, tornou-se igualmente importante. Ainda assim, nas redações – onde as tradicionais práticas e normas da profissão estavam enraizadas – o jornalismo de precisão foi encarado apenas como uma prática secundária, que auxiliava os jornalistas, especialmente os de investigação, a verificar factos para produzir histórias melhores. Não era visto como um novo modo de ver o mundo e de contar o que nele se passa. E enquanto matéria prima, os dados eram encarados como um complemento e não como a matéria principal. As informações recolhidas através de entrevistas a especialistas e da observação direta dos acontecimentos – técnicas tradicionais de recolha de informação – continuavam a estar no pedestal informativo (Coddington, 2015).

Passo a passo, o jornalismo foi timidamente conhecendo o poder do mundo dos dados, o qual mostrou que é possível conhecer mais do que é possível ver. Que é possível ir mais longe. Os dados – elementos quantificáveis e manipuláveis – sozinhos são abstratos, nada significam. Mas quando se relacionam entre si podem revelar verdades, onde existe sentido factual (Ramos, Spinelli, & Estevanim, 2015). E estas têm o potencial para ser a principal matéria prima das histórias e não a secundária. Novas informações surgem da análise de um mar de dados, que de outro modo não seriam reveladas. Novas perspetivas sobre os temas se formam. Os jornalistas deixam de estar dependentes das informações divulgadas por fontes oficiais – através de comunicados de imprensa, por exemplo – e passam a ter a possibilidade de validá-las e até de contestá-las (Bounegru, 2012). O jornalismo, enquanto prática contemporânea, renova-se, viajando no universo dos dados.

Ano após ano, a tecnologia digital à disposição dos jornalistas tornou-se cada vez mais sofisticada. A sua capacidade para processar e armazenar dados e informação foi melhorada e tornou-se possível gerir uma imensidão de dados através de sistemas como o *MySQL* (Teixeira & Junior, 2012). Na passagem do milénio, um novo meio de difusão de informação, a Internet, ganhou força entre os meios de comunicação social a nível mundial (Bastos, 2007). As notícias passaram a ser introduzidas no mundo digital. E uma nova linguagem jornalística, adaptada a este meio, surgiu (Canavilhas, 2006). Na mesma



época, ergueu-se o movimento *open source*, que revolucionou o acesso a dados, informações, bem como a *software*, pois passaram a estar facilmente disponíveis ao público. O direito universal à informação e ao conhecimento começou a ser reconhecido. E as organizações governamentais e não-governamentais passaram a divulgar periodicamente relatórios, documentos e bases de dados sobre a sua atividade. Passaram, portanto, a apresentar traços de transparência perante o público (Howard, 2014). O número de fontes de informação cresceu. E os dados tornaram-se mais acessíveis do que nunca. Passaram, aliás, a ser produzidos em larga escala. A implementação de sistemas e tecnologias de informação no seio empresarial e organizacional – que possibilitou o controlo em tempo real das atividades através de sistemas integrados de gestão, como por exemplo o SAP – e o êxito mundial das redes sociais, acedidas via Internet, são duas grandes nascentes, destacadas por Howard (2014), que germinam biliões de dados todos os dias, todas as horas, todos os minutos. Este fenómeno, em que as ações digitais viram números, não tem data prevista para acabar. A Era do *Big Data* veio, na verdade, para ficar (Pavlik, 2015). E no momento em que tudo – ou quase tudo – é passível de ser descrito através de dados, os jornalistas e as organizações mediáticas enfrentam um grande desafio: dominá-los e interpretá-los, para que seja possível compreender o mundo contemporâneo (Bradshaw, 2012).

É neste contexto de pura evolução em diferentes frentes que os jornalistas passam a ter fácil acesso a bases de dados disponibilizadas e atualizadas pelas organizações, nas quais podem encontrar novas informações. Passam a ter acesso a novas ferramentas tecnológicas que auxiliam e agilizam o processo de análise de dados (Howard, 2014). Passam a ter a possibilidade de realizar trabalhos colaborativos com outros meios de comunicação – como aconteceu no caso *Wikileaks* – e até com o público, através de iniciativas de *crowdsourcing*, um processo em que o público é convidado a analisar documentos e a dar o seu parecer sobre os temas, por exemplo (Baak, 2011). Passam, também, a reconhecer que há valor jornalístico a ser descoberto nos dados, e que apenas é possível encontrá-lo analisando-os. E tudo de floresce de forma natural: as novas técnicas de recolha, análise e apresentação de dados inserem-se num fluxo de trabalho bem definido, num modo particular de ver e de fazer jornalismo, o qual é designado por jornalismo de dados (Bounegru, 2012). Neste campo jornalístico, os dados passam a ter um papel central: é em torno deles que se desenha toda uma história. As metodologias para analisá-los não são uma prática secundária, mas a principal. As técnicas de redação

continuam a ser importantes para transmitir o que há de novo, o que foi descoberto numa avalanche de dados e o seu contexto. E as visualizações ganham um novo encanto, sobretudo no meio digital, pois é a partir delas que o público consegue explorar os dados e encontrar significado neles (Coddington, 2015).

O poder que o fenómeno *Big Data* – que se traduz numa complexa geração de biliões de dados diariamente (Howard, 2014) – e que o mundo tecnológico exerce sobre o jornalismo ainda é maior. Ainda vai mais longe. As linguagens de programação não tardaram a ser aplicadas nos meios de comunicação, quer na construção de *websites* à medida das suas necessidades (através da linguagem HTML, por exemplo), quer na criação de aplicações de notícias – as quais vão ser aprofundadas mais à frente. Naturalmente enraizado nas ciências da computação, o denominado jornalismo computacional tem, segundo Coddington (2015), dois aspetos que o caracterizam: a abstração e a automação. Primeiro, o pensamento abstrato permite quebrar a informação, os problemas, e enquadrá-los num contexto material. E depois, o processo de automação trata de traduzir toda a informação em linguagem de programação, da qual resulta uma aplicação ao serviço do jornalismo. Um projeto de sucesso resultado da colaboração entre jornalistas e especialistas das ciências da computação é o denominado *DocumentCloud.org*. Foi em 2008 que um grupo de profissionais do *ProPublica* e do *New York Times* embarcaram nesta aventura, na qual desafiavam jornalistas e o público a partilhar na Internet – no espaço virtual, a denominada *cloud* – informações, documentos e dados estruturados de interesse público (Cohen, Li, Yang, & Yu, 2011). Outro exemplo são os algoritmos, que sozinhos podem priorizar, classificar e filtrar a informação, auxiliando o trabalho do jornalista. E podem, até, escrever histórias automaticamente. O jornal estadunidense *Los Angeles Times* foi o primeiro a publicar uma notícia escrita por um algoritmo, em 2014. Hoje todas – ou quase todas – as notícias sobre crimes divulgadas pelo mesmo jornal são escrita pelo algoritmo “*Quakebot*” (Howard, 2014). Apesar de produzirem notícias de forma rápida, estes algoritmos estão ainda longe de substituir o trabalho dos jornalistas, pois não têm a perceção sensorial dos acontecimentos da mesma forma que um ser humano. Não têm – ainda – a capacidade de realizar entrevistas a especialistas e também não conseguem refletir sobre os critérios de noticiabilidade relativos aos vários aspetos da história (Howard, 2014). Têm, portanto, limitações que carecem de supervisão humana. Posto isto, promover um trabalho híbrido e colaborativo entre algoritmos e jornalistas, poderá, segundo Lewis (2015), ser um

caminho a seguir. Os algoritmos, por um lado, têm uma capacidade de armazenamento e processamento de informação incomparável. E, por outro, os jornalistas possuem um modo de ver, de narrar e de confirmar a informação sem igual. As qualidades de um preenchem as limitações do outro. E os dois, trabalhando lado a lado, podem complementar-se (Howard, 2014).

O jornalismo abraçou os dados desde cedo. E os modos de recolher, analisar e apresentá-los foram evoluindo com o tempo, que foi trazendo – e ainda traz – novidades no campo tecnológico. Primeiro, a força de vontade venceu as dificuldades e os jornalistas recolhiam e analisavam os dados com recurso a papel e caneta. Depois, os computadores digitais vieram facilitar o processo e a Reportagem Assistida por Computador nasce. Anos mais tarde, Philip Meyer propõe descrever o processo com maior pormenor e esta prática jornalística baseada em dados ganha uma nova vida e um novo nome, o jornalismo de precisão. E recentemente, a revolução digital – marcada pelo aparecimento da Internet, pelo desenvolvimento das tecnologias e pelo movimento *open source* – fizeram respirar o jornalismo de dados (Mogoş, 2016) e o jornalismo computacional (Cohen et al., 2011). Estas abordagens jornalísticas são distintas, mas têm um traço comum: abraçam os dados, para neles encontrar informações novas, de interesse público, que possam mudar – talvez – a vida de quem as lê, ouve e escuta, de quem as explora. É nestas abordagens que se encontram as raízes, a história, as motivações que moldaram o jornalismo que hoje dos dados traz novas histórias (Coddington, 2015).

### **1.3.1 O que o Jornalismo de Dados traz de novo**

O ambiente em que o jornalismo de dados ganha força nas redações é marcado pela revolução digital da informação, que, segundo realça Mogoş (2016), “*mudou a forma como a sociedade contemporânea se relaciona com os dados*”. E, por esse motivo, a forma como é definido e caracterizado, no seio profissional e académico, acaba por ter traços do universo digital que o fez crescer, os quais o distinguem das outras formas de fazer jornalismo a partir de dados.

No sentido lato, o jornalismo de dados é, nada mais nada menos, do que um modo de fazer jornalismo que utiliza dados para contar histórias, tal como refere Henninger (2013). Mas, na verdade, é muito mais do que isso. É uma prática disruptiva que, segundo Lewis (2015), quebra as rotinas tradicionais dos jornalistas para descodificar o lado oculto do mundo social, que está escondido nos dados. É uma prática que, segundo Bounegru

(2012), representa o processo de democratização de recursos (bases de dados e documentos), ferramentas (*software*), técnicas e métodos de análise que outrora apenas estavam disponíveis a especialistas, como jornalistas de investigação, analistas e profissionais de estatística. É, portanto, uma prática híbrida que conjuga dois mundos – o do jornalismo, com o das mais recentes tecnologias (Coddington, 2015) –, que veio melhorar a descoberta e verificação de factos e, assim, potenciar o papel do jornalista enquanto guardião da verdade (Howard, 2014).

O potencial do jornalismo de dados começou a ser percebido quando o caso *Wikileaks* surgiu, na primeira década do milénio. A dificuldade em analisar cerca de 92 mil documentos vazados do exército dos Estados Unidos da América sobre a Guerra do Afeganistão foi tal que os meios de comunicação internacionais uniram-se. O jornal britânico *The Guardian*, o estadunidense *The New York Times* e o alemão *Der Spiegel* são alguns exemplos. Além disso, foi também necessário encontrar um fluxo de trabalho que permitisse organizar e analisar tamanha quantidade de dados e informação: o jornalismo de dados. Esta estratégia adotada pelas redações levou-as a descobrir, por exemplo, que milhares de civis afegãos morreram pela mão de militares norte-americanos. Levou-as a descobrir uma realidade até então desconhecida, que merecia ser partilhada com o mundo. Através de visualizações interativas, estes jornais mostraram ao público quando e onde as mortes ocorreram e até em que circunstâncias – o qual apenas foi possível através da utilização das ferramentas *open source*, como o *Google Fusion Tables*. Foi neste contexto que o valor do processo do jornalismo de dados foi descoberto (Baak, 2011). O caso *Wikileaks* é mesmo considerado por Baak (2011) uma força motriz que impulsionou o desenvolvimento do jornalismo de dados, pois foi este que mostrou ao mundo que é possível encontrar verdades e revelar informações de forma clara e compreensível ao público mergulhando num mar de dados.

As redações pelo mundo fora despertaram para o poder do jornalismo de dados e, por conseguinte, passaram a empregá-lo (Coddington, 2015). O diário britânico *The Guardian* foi um dos primeiros a criar uma rubrica *online* dedicada ao jornalismo de dados, o denominado *DataBlog*, que logo começou a dar especial destaque às visualizações, as quais permitem explorar os dados (Mogoş, 2016). Nos seus primeiros anos de vida – entre 2009 e 2010 –, o *DataBlog* lança um projeto no qual desafia os seus leitores a colaborar na análise de 460 mil páginas dos relatórios que mostravam as despesas dos membros do Parlamento do Reino Unido. O objetivo passava por sinalizar

e classificar aspetos questionáveis. A colaboração do público foi notável e a história redigida a partir dos resultados levou mesmo à revisão das despesas por parte do Governo britânico da altura. Esta iniciativa do diário britânico é, para Coddington (2015), um exemplo claro do que é fazer jornalismo de dados. Focar os dados disponíveis a todos, os denominados *open-data*. Solicitar a colaboração do público para analisá-los, através de iniciativas de *crowdsourcing*. Refletir, examinar e retirar conclusões dos resultados da análise da audiência. Escrever histórias de interesse público inserindo nelas visualizações, que promovem a exploração e compreensão do conteúdo. E, assim, relevar verdades desconhecidas que têm o potencial de mudar o mundo.

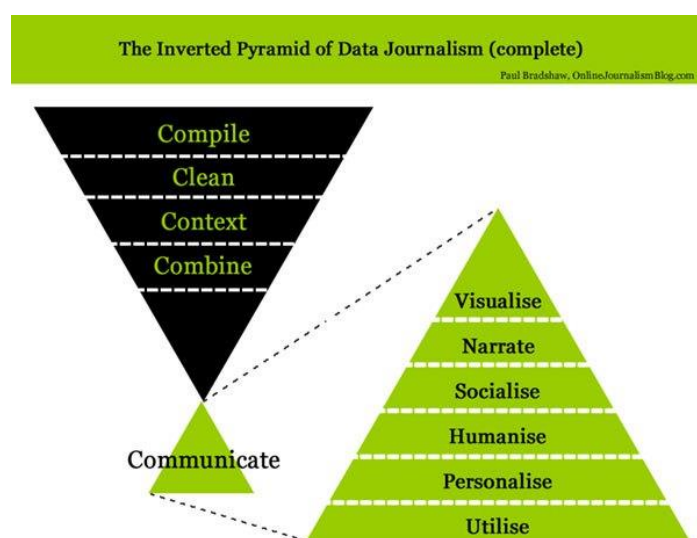
O jornalismo de dados passou a ser uma prática cada vez mais presente nos meios de comunicação mundiais a partir da segunda década do século XXI. O *New York Times* lança, em 2014, o projeto que prometeu apresentar notícias baseadas em dados de uma outra forma, o denominado *The Upshot*. Reconhecendo o poder dos dados no ambiente digital para clarificar o mundo, o diário norte-americano aposta num formato onde é possível navegar nos conteúdos informativos e explorar gráficos e outras visualizações interativas. No mesmo ano, as redações norte-americanas do *Washington Post* e do *Wall Street Journal* lançam também projetos que dão especial destaque aos dados e que procuram tirar partido do potencial das visualizações, as quais permitem compilar dados complexos transformando-os em informação compreensível e visualmente atrativa (Mogoş, 2016). Em Portugal, o semanário Expresso lança em 2015 uma rubrica audiovisual dedicada ao jornalismo de dados: o 2:59, no qual se propõe a explicar o mundo em dois minutos e 59 segundos partindo de diversos dados (Expresso, 2016).

Estes foram os primeiros meios de comunicação do mundo a criar projetos digitais onde o jornalismo de dados é protagonista. Para que todos estes – e outros – projetos jornalísticos baseados em dados fossem possíveis de materializar, um novo e particular fluxo de trabalho foi implementado no seio dos meios de comunicação social, o do jornalismo de dados. E a sua integração gerou um sismo de mudanças nas redações, nas rotinas produtivas e nos próprios jornalistas. No epicentro estão as novas tecnologias, em particular os sistemas de informação, que se tornaram as ferramentas chave de todo o processo: desde a recolha, limpeza e análise de dados até à criação de visualizações interativas. O olhar jornalístico não é descurado. Saber identificar histórias nos dados e saber contá-las continua a ser importante. Mas dominar os sistemas de informação tornou-se crucial para construí-las.

### 1.3.2 O Jornalismo de Dados enquanto fluxo de trabalho

Enquanto prática, o jornalismo de dados é definido como um processo, um fluxo de trabalho orientado onde há recolha e seleção de bases de dados. Onde há limpeza e análise dos mesmos para que seja possível encontrar tendências e anomalias que caracterizam o mundo social dos nossos dias. Onde as informações novas, com especial interesse público, são colocadas num contexto. Onde há, portanto, uma história a ser contada, não só com recurso a texto como também a visualizações, como, por exemplo, gráficos, diagramas, mapas e infografias (Coddington, 2015).

O jornalista britânico Bradshaw (2011) propõe dar uma nova vida à tradicional pirâmide invertida – anteriormente caracterizada – ajustando-a ao jornalismo de dados. Esta nova superestrutura procura descrever o processo de jornalismo de dados em cinco fases distintas, mas complementares, tal como ilustrada a Figura 1.1: a fase de compilação de dados, a que trata a limpeza dos mesmos, a que os contextualiza, a que combina diferentes bases de dados e, por último, a referente à comunicação da história.



**Figura 1.1** - A Pirâmide Invertida do Jornalismo de Dados

**Fonte:** Cópia Integral de Bradshaw (2011b)

Na visão de Bradshaw (2011b), tudo começa com uma questão: ou o jornalista tem uma pergunta que apenas pode ser respondida com recurso a dados, ou existe um conjunto de dados que devem ser questionados. Perante este ponto de partida, há um primeiro caminho a percorrer, o da compilação de dados, que podem ser recolhidos de diferentes formas. Podem ser descarregados das páginas *web* de organizações governamentais e não governamentais. Podem ser recolhidos através da análise de documentos e relatórios técnicos. E também podem ser reunidos por meio de pesquisas, observações e projetos de

*crowdsourcing*. Em segundo plano, há que limpar o conjunto de dados recolhidos para garantir a qualidade e confiança dos mesmos. Bradshaw (2011b) considera que, em geral, há dois modos de limpar os dados: o primeiro consiste em remover o erro humano e o segundo em converter os dados em formatos consistentes e homogêneos. A formatação em texto (em formato PDF, por exemplo), a existência de células vazias e a presença de diferentes formas para referir uma entidade, uma data ou um local são aspetos comuns que carecem de tratamento. Ferramentas como o *Microsoft Office Excel* e o *Google Refine* permitem localizar e substituir dados, usar filtros e classificá-los de forma simples, o que auxilia os jornalistas neste processo. Na terceira fase, Bradshaw (2011b) destaca a importância de conhecer o contexto da origem dos dados – quando foram recolhidos, por quem e com que propósito – para que seja possível validar a sua confiabilidade. O quarto aspeto diz respeito à possibilidade ou necessidade de combinar mais do que uma base de dados para contar uma história. Apresentar a distribuição de dados no espaço é um exemplo de referência desta fase de combinação, segundo Bradshaw (2011b). Através da utilização de aplicações *open source* da *Google*, como o *Google Fusion Tables* e o *Google Maps*, os jornalistas podem agregar bases de dados e criar mapas interativos. Por último, surge a fase de comunicar, a fase em que é produzida a história final e em que é delineada uma estratégia para a transmitir ao público de forma eficaz.

Bradshaw (2011a) aprofunda esta última etapa em seis níveis diferentes. Primeiro destaca as visualizações – como mapas, gráficos, infografias e animações – que tornam as histórias mais ricas e a sua comunicação efetiva. As visualizações possuem uma real força de atrair e transmitir informação num instante, com apenas um olhar. Mas esta vantagem pode tornar-se numa fraqueza se não for delineada uma estratégia para promover o envolvimento e a exploração do conteúdo, através de hiperligações, por exemplo. Um outro nível da comunicação é a narrativa, a qual procura contextualizar os números, relacionando e interpretando-os. O jornalista britânico destaca também a importância de envolver a esfera social no processo de jornalismo de dados – através de estratégias de *crowdsourcing*, por exemplo – para que seja possível conhecê-la melhor. Conhecer de perto a visão do público é um ponto de partida para chegar a bom porto, onde os números têm significado para as pessoas, onde são contadas histórias de humanos para humanos. Este é o quarto nível de comunicação referido por Bradshaw, (2011a), o que reconhece o impacto que uma história tem – ou pode ter – na vida das pessoas. E para que o público consiga explorar a história consoante as suas preferências pessoais, para que consiga,

portanto, personalizá-la com intuito de encontrar significado nas suas vidas, é importante que o conteúdo possua traços de interatividade, o que apenas é possível no ambiente digital. Este nível, o da personalização, leva-nos ao último, no qual o conteúdo é passível de ser utilizado pela audiência, através de aplicações digitais, como por exemplo calculadoras e ferramentas orientadas por *Global Positioning System* (GPS).

Este é o modo como o jornalista britânico Bradshaw (2011b) descreve o jornalismo de dados enquanto fluxo de trabalho, no qual dá especial destaque à recolha e ao tratamento de bases de dados, bem como à forma como os resultados são comunicados ao público. Ainda assim, há uma fase importante, referida com frequência na literatura, que o autor não aprofunda: a da análise de dados. Esta é uma fase com especial relevância, pois é nela que o jornalista aplica técnicas estatísticas e exploratórias de dados para neles encontrar informações relevantes e de interesse público (Henninger, 2013).

#### ***1.3.2.1 Da análise de dados nascem as histórias***

Abordar profundamente a fase de análise de dados não vem ao acaso. Esta é a etapa do processo do jornalismo de dados que surge imediatamente antes da construção do produto final – a história – e que revela dois aspetos fundamentais. Por um lado, evidencia informações escondidas numa imensidão de dados, que acabam por determinar o fio condutor de uma história. E por outro, mostra também os dados mais relevantes que devem ser comunicados ao público através de visualizações.

Quando descreve o processo de jornalismo de dados, Henninger (2013) explica como é possível compreendê-los analisando-os. Analisar bases de dados com um tamanho expressivo é, por si só, uma tarefa complexa, mas o autor lembra que hoje os jornalistas têm à sua disposição os mais recentes aplicativos e sistemas de informação que facilitam, cada vez mais, o processo. Um deles é o *Microsoft Office Excel* e, em especial, o seu complemento, o *Power Pivot* (este último, atualmente, integrado na estratégia de *Business Intelligence* da Microsoft num conjunto que se designa por *Power BI*). Henninger (2013) faz referência a este complemento do *Excel* como uma solução ótima para trabalhar enormes quantidades de dados, pois permite selecionar, filtrar, cruzar e calcular diversas variáveis de diferentes modos. E permite ainda criar visualizações adequadas à natureza dos dados – como gráficos, mapas e diagramas –, que facilitam a análise e a compreensão dos mesmos. Deste modo, é possível encontrar padrões, anomalias e relações entre os dados. É possível, portanto, encontrar evidências e descobrir conhecimento mergulhando



neles. E são estas novidades descobertas nos dados que podem originar uma nova história, ou uma nova perspectiva sobre um tema.

Sobre como nascem as histórias baseadas em dados, Rosenbaum (2012), jornalista da BBC, vai mais longe. O autor defende que as técnicas utilizadas para analisar os dados influenciam diretamente a forma como as histórias são contadas. Usar métodos de medição, como a contabilização ou a totalização de uma variável, leva o jornalista a obter resultados quantitativos que indicam, por exemplo, a frequência absoluta ou relativa de ocorrências num determinado espaço e tempo. Para adicionar contexto, o jornalista pode – através de um gráfico – explorar a evolução ao longo do tempo de uma variável e perceber quando é que atingiu o seu máximo e mínimo num determinado período de tempo, e quando obteve o maior crescimento. Pode ainda proceder a uma análise comparativa externa, na qual explora dados que medem a mesma realidade em vários países, por exemplo. Pode aplicar técnicas estatísticas para correlacionar duas variáveis, com o intuito de perceber se há ou não uma associação forte entre elas.

Tendo em conta a relação existente entre as técnicas de análise e as narrativas que delas resultam, Rosenbaum (2012) propõe caracterizar diferentes tipos de histórias baseadas em dados, como as histórias mensuráveis, que resultam de uma medição, ou as associativas que relacionam positivamente duas variáveis. Mas, na verdade, uma só base de dados pode ser explorada em diferentes perspetivas, utilizando, para tal, diferentes técnicas. E a história é, nada mais nada menos, do que um conjunto orientado de informações descobertas a partir da aplicação de todas essas técnicas. Não é, portanto, uma história que resulta só da medição de uma variável ou apenas de uma comparação. E, por esse motivo, na perspetiva da investigadora, não faz sentido classificar as histórias do modo sugerido por Rosenbaum (2012). Uma história que resulta do jornalismo de dados é – ou pode ser – um conjunto de análises, de perspetivas, de técnicas, e não apenas uma.

Uma parte que integra – ou que pode integrar – uma história que dos dados germina é a apresentação de dados através de visualizações. Esta é, segundo Mogoş (2016), a forma mais clara, precisa e eficiente de transmitir informação complexa ao público. Apresentar dados técnicos, que geralmente são passíveis de serem explorados em vários níveis de detalhe, é um exemplo (Henninger, 2013). Ainda assim, para que a informação contida nos dados seja transmitida de forma eficaz, Sarah Cohen (2012) recorda que os dados selecionados devem ser precisos e significativos para o público. Cumprindo estes

requisitos, as visualizações podem ser uma importante arma para os meios de comunicação, pois têm o poder de cativar o público, o qual possui uma natural atração por conteúdos visuais (Cohen, 2012).

Para que uma visualização cumpra o seu principal objetivo de informar, de convidar o público a conhecer mais sobre um determinado assunto, Henninger (2013) sublinha a importância de, numa primeira instância, examinar os dados para, em seguida, compreender qual o tipo de visualização que mais se adequa aos mesmos. Para os dados que revelam a evolução ao longo do tempo, o tipo de visualização mais adequado, na perspectiva do autor, será o gráfico linear. Para variáveis categorizadas, a melhor opção será um gráfico de barras. Para variáveis quantitativas – que expressam uma frequência absoluta ou relativa de uma variável –, os histogramas são os mais apropriados. E para distribuir valores no espaço, os mapas são os eleitos. Neste último tipo – os mapas –, o autor destaca ainda diferentes formas de representar a informação: em mapas de pontos, que mostram a distribuição espacial de uma variável, evidenciando a sua densidade e quantidade; em cartogramas, onde a dimensão de uma variável é representada por símbolos – tendo em conta uma escala – e localizada no espaço; em “*choropleth maps*”, que usam o código de cor para agregar dados por áreas geográficas; e em mapas de calor, onde os dados – representados por gradientes de cor – são introduzidos numa matriz de duas dimensões (Henninger, 2013). Para desenvolver estes e outros tipos de visualizações, os jornalistas têm hoje várias ferramentas à sua disposição. Algumas permitem até criar visualizações interativas no meio digital, que convidam o público a explorar os dados de forma dinâmica (Mogoş, 2016). Uma das ferramentas destacadas por Henninger (2013) é o *Tableau Software*, usado em jornais como *The Guardian*, *The New York Times* e o *Expresso*.

São várias as formas de apresentar os dados através de visualizações, mas também são vários os erros que os jornalistas podem cometer ao fazê-lo. Colocar demasiada informação técnica pode, por exemplo, tornar os gráficos confusos e pouco apelativos. Ainda assim, a falta de informação, quando relevante – como a origem dos dados, a localização no espaço e no tempo –, também é considerada uma falha, pois pode levar os leitores a desconfiar da sua credibilidade. O uso exagerado de detalhe pode também tornar uma visualização mais complexa do que clara. Estes são apenas alguns exemplos, destacados por Mogoş (2016), dos erros que se podem cometer quando se desenvolve uma visualização. Ainda assim, o maior dos erros, segundo Cohen (2012), é não saber

identificar situações em que incorporar visualizações na história não é adequado. Quando, por exemplo, um gráfico não consegue ter o mesmo impacto que uma narrativa para transmitir uma mensagem. Isto é, quando há um outro meio capaz de contar a história com maior qualidade e eficácia do que uma visualização. Quando há poucos dados a serem explorados, pois dificulta a contextualização dos mesmos. Ou quando os dados não expressam uma tendência clara, devido à sua fraca variabilidade (Cohen, 2012). Em resultado, as visualizações de dados podem ter exatamente o efeito contrário do pretendido: podem confundir ao invés de clarificar o público, podem afastá-lo em vez de aproximá-lo (Mogoş, 2016).

Cada história é uma história, pelo que deve ser encarada com a devida singularidade. Os dados que a originam devem ser tratados e analisados de um modo particular. E a forma como a informação é transmitida deverá ser adaptada à sua natureza. O importante é que o produto final, a história baseada em dados, possua um *design* informativo adequado. Isto quer dizer que a convergência de meios que integram a história – como fotografias, texto, gráficos, vídeos, áudios, infografias, entre outros – deve encontrar uma harmonia, uma sintonia para que seja possível contar ao público um conjunto de verdades de forma clara e transparente e, sobretudo, entendível (Henninger, 2013). Há várias formas de alcançar a referida harmonia. E uma delas é desenvolvendo aplicações de notícias (as denominadas *news apps*, na língua inglesa) no seio digital. Esta estratégia já é adotada por vários meios de comunicação, como é o caso do diário britânico *The Guardian* e dos jornais dos Estados Unidos da América (EUA): *The New York Times* e o *ProPublica* (Howard, 2014).

### ***1.3.2.2 Aplicações de notícias: uma nova forma de contar histórias***

Para construir uma aplicação no seio jornalístico, primeiro há que focar o seu destinatário, o público, bem como as suas necessidades. Há que questionar quais são as informações que podem ser úteis para as suas vidas (Davis, 2012). Este foi o exercício que os jornalistas do *ProPublica* fizeram em 2016, quando criaram uma aplicação inovadora, que permitia pesquisar informações sobre a segurança das clínicas norte-americanas de diálise renal.

Na origem da aplicação estão as bases de dados, as quais devem possuir duas características importantes, segundo Klein (2012), jornalista e programador do *ProPublica*. Por um lado, devem ser abrangentes o suficiente de modo a incluir informações sobre toda uma nação. E, por outro, devem também incluir detalhes importantes sobre um tema, passíveis de serem explorados. Os dados são, portanto, a principal matéria prima das aplicações, não descurando a manchete, a entrada, e as narrativas que a mesmas integram, ou podem integrar. Ao nelas navegar, o leitor vai-se deparando com diferentes dados sobre uma realidade nacional, que surgem consoante a sua pesquisa. Deste modo, um leitor consegue não só ter uma nova perspetiva sobre um acontecimento nacional como também descobrir uma ligação pessoal com o mesmo (Klein, 2012). Esta forma interativa de apresentar os dados e contar histórias é, na visão de Howard (2014), uma solução ideal para o leitor compreender o conteúdo em profundidade e para nele encontrar utilidade para a sua vida. E é também uma forte alternativa para chegar até às gerações mais jovens, que vivem num mundo marcado pela utilização diária de aparelhos móveis ligados à Internet.

*“Na melhor das hipóteses, as aplicações de notícias não contam apenas uma história, contam a história de quem lê”*, sublinha Howard (2014). O verdadeiro desafio, segundo Davis (2012), é agregar valor a estas aplicações de forma a que este não se dissipe com o tempo, que não seja perecível como é o valor das notícias.



## **2 A integração do Jornalismo de Dados e as mudanças subjacentes**

### **2.1 As mudanças nas redações**

Integrar o jornalismo de dados de forma efetiva no seio das redações significa adaptar as rotinas produtivas dos jornalistas a um novo fluxo de trabalho. Significa investir em sistemas de informação adequados para produzir conteúdos que possam proporcionar aos leitores uma autêntica viagem multimédia entre os dados. Significa – ou pode significar – também recrutar profissionais de áreas distintas. E requer, sobretudo, pensar num modelo de negócio de forma a tirar partido das potencialidades desta tendência jornalística internacional.

Hoje, as redações abraçam as novas ferramentas tecnológicas de tal modo, que estas passam a apoiar quase todos – senão todos – os processos jornalísticos. Posto isto, integrar o jornalismo de dados nas redações significa, por um lado, dar novas utilidades a algumas destas ferramentas, empregando-as com novos propósitos. E, por outro, significa também investir em novos sistemas de informação. O processo criativo de narrativas baseadas em dados poderá ser apoiado pelo sistema de processamento de texto já integrado, como por exemplo o *Microsoft Office Word*. E a análise de dados suportada pelo *Microsoft Office Excel* (Henninger, 2013). Caso a história inclua imagens e vídeos, os mesmos também podem ser editados nos respetivos sistemas já existentes nas redações. Como, por exemplo, o *Adobe Photoshop*, para conteúdos fotográficos, e o *Adobe Premiere*, para conteúdos audiovisuais. Mas para criar visualizações de dados – sobretudo interativas – geralmente deve-se apostar em novos sistemas de *design* gráfico, como é o caso do *Tableau Software* (Henninger, 2013).

Usar estes sistemas em prol do jornalismo de dados significa – ou pode significar – também recrutar novos profissionais capazes de os manusear, como é o caso de *designers* de comunicação visual e infografistas. E para criar aplicações de notícias, onde é possível explorar o seu conteúdo de forma interativa, é necessário integrar profissionais que dominem linguagens de programação. As equipas tornam-se multidisciplinares: os jornalistas passam a trabalhar em parceria com profissionais de outras áreas. Para construir histórias baseadas em dados sobre as temáticas que mais cativam o público, criar equipas destinadas a analisar os dados gerados *online* – quer nas páginas *web* dos *media*, quer nas suas páginas nas redes sociais – torna-se igualmente importante. Conhecer, por exemplo, os temas que geram mais visualizações e reações e, portanto, os assuntos que

mais atraem o público, pode ser importante na hora de decidir quais os temas que devem ser aprofundados. Este é apenas um dos vários raciocínios possíveis, que mostra como é possível tirar partido dos milhões de dados gerados todos os dias pelos leitores. Estes dados sobre o público – gerados num contexto digital – são valiosos, pois permitem conhecê-lo em profundidade, permitem conhecer de perto as suas preferências, o seu percurso de leitura. E, por isso, estes dados tornam-se importantes para delinear estratégias de negócio focadas nos interesses do público (Howard, 2014).

Para Howard (2014), conhecer a audiência é apenas o primeiro passo para construir um modelo de negócio sustentável para os meios de comunicação social, pois é com base nessas informações que é possível desenhar uma proposta de valor para o público. Transformar dados complexos – que permitem conhecer realidades distintas – em informação facilmente compreensível é o passo seguinte. Introduzi-la em aplicações interativas – anteriormente caracterizadas – pode ser uma opção de apresentação de informação atrativa e eficaz, segundo o autor. O importante, de acordo com Lorenz (2012), é dar a possibilidade ao público de se relacionar com os biliões de números que todos os dias são gerados. É criar histórias onde os leitores têm a possibilidade de explorar os dados para neles encontrar significado para as suas vidas. É, segundo Lorenz (2012), criar uma nova proposta de valor para os leitores. E, na perspetiva do autor, publicar dados credíveis que colocam em perspetiva questões complexas sobre as mais diversas áreas de interesse público – como a educação, a saúde, o ambiente, o emprego, os direitos humanos, entre outras – poderá ser a estratégia a seguir para reanimar a confiança do público no jornalismo. Mas até que ponto poderá o jornalismo de dados melhorar o ambiente de desconfiança no jornalismo?

Na verdade, já são vários os meios de comunicação social a adotar estratégias de negócio assentes no jornalismo de dados, sobretudo nos EUA, as quais, segundo Trasel (2014), procuram recuperar a confiança do público. O caso da *Bloomberg* é um exemplo destacado tanto por Howard (2014) como por Lorenz (2012), que espelha o sucesso que uma agência de notícias pode ter quando investe na divulgação de notícias baseadas em dados – neste caso em particular, em dados financeiros. Um outro caso de sucesso é o da *Texas Tribune*, um meio de comunicação social sem fins lucrativos (Howard, 2014). O seu investimento em aplicações de notícias, gráficos e bases de dados foi tal, que hoje já conta com mais de 200 conteúdos deste tipo, os quais estão inseridos numa secção com especial destaque na sua página *web*. Além do referido, a sua estratégia de negócio está

assente em várias fontes de receitas. Howard (2014) refere que para além da publicidade, a sustentabilidade financeira do *Texas Tribune* é também assegurada por doações, rendas ganhas, subscrições e eventos. O caso do *ProPublica*, um meio de comunicação social sem fins lucrativos dos EUA, também não passa despercebido. Investe na *DataStore*, na qual publica bases de dados brutos – isto é, dados antes de serem analisados – que tanto podem ser gratuitas como pagas. E também constroem aplicações tendo por base as suas próprias análises dos dados.

Estes são apenas alguns meios de comunicação social que mostram que é possível beneficiar das mudanças que a revolução digital provoca nas redações. Com o trabalho conjunto entre profissionais de diferentes áreas torna-se possível edificar produtos jornalísticos diferentes, mais úteis e apelativos para todos. Mostram também que é possível estruturar modelos de negócio de sucesso baseados em diferentes projetos digitais, nos quais os dados possuem um papel central. Estes são apenas alguns meios de comunicação que mostram o potencial que os dados podem ter na construção de um jornalismo digital mais credível aos olhos do público. E que investir neste sentido pode ser o caminho para reconstruir a sua confiança (Pavlik, 2014).



## **2.2 Um novo perfil emergente: o jornalista de dados**

O modo de olhar e de narrar próprio do jornalista é desafiado quando se confronta com uma avalanche de dados. No universo do jornalismo de dados não basta dominar a arte de bem escrever e ter a curiosidade pela descoberta de novas histórias. É preciso conhecer um vasto leque de técnicas de análise de dados para conseguir compreendê-los, para conseguir traduzir o seu significado. A literacia estatística torna-se fundamental, bem como o uso ágil das mais recentes tecnologias e sistemas de informação, capazes de analisar dados e desenhar gráficos e infografias interativas. Saber programar aplicações torna-se igualmente relevante, assim como saber conduzir iniciativas de *crowdsourcing*. São vários os saberes e as capacidades que um jornalista deve cultivar. E estes, segundo López-García, Toural, & Rodríguez-Vázquez (2016) caracterizam um novo perfil emergente: o jornalista de dados.

Conhecer a natureza dos dados, saber como cruzar diferentes bases de dados, como interpretá-las e ter a habilidade de questionar e de refletir sobre a sua veracidade são capacidades, destacadas por Kayser-Bril (2012), que um jornalista de dados deve desenvolver. De acordo com o autor, há três questões essenciais que um jornalista de dados deve colocar para cultivar estas habilidades. Primeiro, o jornalista deve ter espírito crítico quanto à origem dos dados para averiguar a sua credibilidade. E deve questionar também as informações obtidas a partir da sua análise. Isto quer dizer que é importante recolher esclarecimentos junto de especialistas para verificar a veracidade dos resultados obtidos, assim como para justificá-los. Depois, o jornalista deve questionar-se sobre o que é possível aprender com os dados. Neste campo, dominar a literacia estatística é crucial. Kayser-Bril (2012) refere que calcular a média, a mediana e a moda, assim como a frequência absoluta e relativa de uma variável podem ser pontos de partida para compreender os dados. E depois há várias formas de comunicá-los ao público de forma simples. A título de exemplo, o autor explica que uma frequência natural (1 em cada 100) é uma estratégia mais assertiva e entendível aos olhos do público do que uma frequência relativa (1%). A tradução dos resultados obtidos em informação facilmente perceptível pela audiência é, portanto, um grande desafio para os jornalistas de dados. A terceira questão mencionada por Kayser-Bril (2012) diz respeito ao nível confiança da informação obtida. Será que uma amostra de mil indivíduos é capaz de traduzir uma realidade que abrange um universo de milhões de pessoas? Para responder a esta questão, o autor destaca novamente a importância de conhecer conceitos e cálculos estatísticos. Neste caso

concreto, a margem de erro, que indica o intervalo de erro dos resultados baseados na amostra, e o nível de confiança, o qual representa a probabilidade de obter o mesmo intervalo de confiança se a recolha de dados for direcionada a outras amostras com as mesmas características (Agranonik & Hirakata, 2011). Portanto, se amostra tiver mais de mil indivíduos é possível obter uma margem de erro inferior a 3%, para um nível de confiança de 95%. Isto significa que o intervalo de confiança dos resultados obtidos terá três pontos percentuais acima e abaixo da sua distribuição real pela população. Posto isto – e respondendo à questão supracitada –, os resultados obtidos partindo de uma amostra de mil pessoas podem representar uma realidade que abrange uma população de milhões de pessoas (Kayser-Bril, 2012). Mas há que ter em conta que a margem de erro pode sempre diminuir, se a amostra aumentar (Agranonik & Hirakata, 2011).

Conhecer os referidos conceitos estatísticos e dominar os sistemas de cálculo são dois saberes que dão força aos jornalistas para enfrentar os vários desafios que vão surgindo durante o processo: desde a recolha de dados à redação da história. Se o formato dos dados for em PDF, por exemplo, o jornalista deverá saber como os copiar, importar ou converter numa folha de cálculo de forma a não perder informação. Aquando o cruzamento de bases de dados com origens distintas, o jornalista de dados deverá ter a perícia de perceber como reduzir as variáveis a analisar às mesmas unidades de medida. E perante as pressões de tempo, os jornalistas deverão ter destreza analítica para perceber quais são os aspetos mais relevantes a analisar e a comunicar ao público (Appelgren & Nygren, 2014).

Face à proliferação de dados no ambiente digital a todo o momento, este perfil, o do jornalismo de dados, está a ser cada vez mais procurado pelos meios de comunicação social a nível internacional, como pelo *The Guardian*, *The New York Times* e *ProPublica* (López-García et al., 2016). E face a esta necessidade das redações, já há instituições de ensino a integrar programas de estudo para jornalistas nas áreas de ciências de computação, da estatística e da gestão de bases de dados. Um deles é o *Lede Program*, que resulta de uma parceria entre o departamento de ciência de computação e o de jornalismo na Universidade da Colômbia. Este programa foi desenhado para familiarizar os estudantes de jornalismo com os dados, os códigos, os algoritmos, os conceitos que se estão a tornar cada vez mais relevantes no jornalismo de dados (Kirkpatrick, 2015). Uma outra iniciativa parte da *Google*, que investiu num laboratório de notícias digital no qual oferece vários cursos, como de narração de histórias multimédia, de jornalismo de dados

e de verificação da veracidade de conteúdos. Nestes cursos, os jornalistas têm a possibilidade de aprender, de forma gratuita, a utilizar diversas ferramentas do gigante tecnológico Google, que auxiliam o processo: desde a identificação de temas quentes à criação de visualizações interativas. O *Google Trends*, o *Google Fusion Tables*, o *Google Public Data Explorer*, o *Google Consumer Surveys* e o *Google Crisis Map* são alguns exemplos (Google, 2018).

O mundo do jornalismo está em constante mudança. E num momento em que os dados – gerados a todo o instante – têm o potencial de traduzir realidades com relevante interesse público, torna-se crucial saber analisá-los e compreendê-los. O papel dos jornalistas de dados ganha importância hoje mais do que nunca (López-García et al., 2016). Inculcar os saberes e as capacidades próprias que o caracterizam nas universidades é um primeiro passo para formar profissionais com este perfil, que é cada vez mais procurado pelos *media* internacionais. Mas ter a iniciativa própria de conhecer mais sobre o mundo dos dados, de aprender a manusear as ferramentas adequadas para analisá-los, de saber como traduzir os resultados em informação entendível para o público é um passo ainda maior e que apenas depende dos jovens estudantes de jornalismo e dos próprios jornalistas.

### **2.3 Uma viagem interativa pelos dados: o que muda para o público**

Na atmosfera digital em que hoje viajamos, a comunicação deixa de ter apenas um sentido. Passa a ter dois. O público tem o poder de manifestar a sua opinião. Tem, aliás, hoje mais informação disponível do que alguma vez teve. E face à abundância de informação, o público passou a ter o poder de decisão quanto aos conteúdos que consome. E se paga ou não por eles. Tem o mundo da informação nas suas mãos. Perante este cenário, os *media* veem-se desafiados a criar novos laços com as pessoas, a compreender o que mais as atrai, para criar conteúdos ajustados às suas novas necessidades.

As pessoas rapidamente demonstraram uma predisposição por fazer ouvir a sua voz no seio dos meios de comunicação social: quer nas suas páginas *web*, quer nas páginas oficiais das redes sociais. Redigir comentários e partilhar conteúdos são algumas formas que o público tem de se manifestar. A sua vontade de participar no jornalismo também é evidente. São vários os vídeos e as imagens captadas pela audiência, que mostram uma realidade de interesse público, que hoje fazem parte do conteúdo das notícias. Os cidadãos assumem o papel de jornalista quando têm o privilégio de comunicar um acontecimento relevante em primeira mão (Pavlik, 2014). Estes comportamentos revelam uma tendência natural da audiência para interagir diretamente no mundo jornalístico. E este indicador pode revelar fortes probabilidades de sucesso para iniciativas que cultivam a interatividade.

Antes de produzir um conteúdo jornalístico baseado em dados, Geere (2012) considera que primeiro é preciso conhecer o público a que se destina. É preciso identificar os temas que mais interesse público revelam. Para obter estas informações, analisar as tendências nas redes sociais – como o número de reações, de comentários e partilhas – e as visualizações de conteúdos nas páginas *web* – registado no *Google Analytics*, por exemplo – são dois caminhos possíveis. Em segundo plano, o autor refere a importância de delinear estratégias de apresentação do conteúdo para cativar o público. Neste ponto, Geere (2012) defende que é preciso investir em conteúdos interativos – o que vai ao encontro com a tendência supracitada. Criar visualizações interativas e aplicações de notícias são dois exemplos que despertam a exploração de dados de forma pessoal e, por conseguinte, aceleram a compreensão do significado dos mesmos. Abrir novos canais de comunicação – como espaços para comentários privados – não só estimulam a colaboração do público como também trazem novas ideias e perspetivas sobre os temas. De acordo com Appelgren & Nygren (2014), as iniciativas de *crowdsourcing* são também uma

importante ferramenta que cativa a participação do público, já que solicita a sua colaboração para a interpretação de documentos, integrando-o no processo de produção de conteúdo jornalístico. Além do referido, estas iniciativas são, cada vez mais, uma importante fonte de informação para os jornalistas, que agilizam o processo de interpretação de documentos, geralmente complexos. Um outro ponto importante destacado por Geere (2012) tem que ver com a transparência da informação e dos dados apresentados. Isto quer dizer que publicar as bases de dados utilizadas para contar histórias e, portanto, dar a possibilidade de descarregá-las ou acedê-las via Internet é um passo a seguir para a construção de uma relação de confiança com o público.

Envolver o público nas histórias baseadas em dados é um desafio enorme para os jornalistas. Outro – talvez ainda maior – é melhorar a sua confiança no próprio jornalismo. Mas investir em jornalismo de dados poderá ser um caminho a seguir, já que cria uma série de possibilidades para os leitores. Cria a possibilidade de compreender um conteúdo complexo em profundidade. Cria a possibilidade de interagir com os conteúdos, de navegar entre os dados segundo uma perspetiva pessoal. O jornalismo de dados estimula, portanto, a criação de uma empatia com a história e com o próprio trabalho jornalístico, o que promove – ou pode promover – a criação de valor (Ramos et al., 2015).

### **3 O semanário que se reinventa: o jornal Expresso**

O presente capítulo procura explorar o passado e o presente do semanário Expresso a vários níveis: como surgiu, como se adaptou – e adapta – ao meio digital, como trabalham os seus profissionais e também como integra um projeto digital dedicado ao jornalismo de dados, o 2:59. Para explorar as referidas matérias, a investigadora aplicou diferentes técnicas de recolha de informação. O recurso à literatura foi fundamental, mas as entrevistas e a observação participante permitiram ir mais longe e conhecer de perto a redação, bem os seus profissionais.

Para explorar a história do semanário Expresso e compreender de que forma se moldou ao mundo digital, a exploração do estado da arte e a realização de entrevistas, quer formais quer informais, foram as principais técnicas de recolha de informação aplicadas. Para realizar as entrevistas, onde existe uma interação direta entre dois atores, foi redigido um guião de questões. E como a hipótese de explorar outras questões que surgem com naturalidade não foi descurada, as entrevistas são semiestruturadas (Rodrigues, 2015).

Para melhor conhecer as rotinas produtivas do jornal Expresso, bem como a forma como estes profissionais integram a rubrica 2:59, a observação participante foi fundamental. Ao aplicar esta técnica, que possibilita ao investigador analisar e perceber diretamente uma realidade envolvendo-se nela (Oliveira, 2012), foi possível não só observar diariamente as rotinas dos jornalistas do Expresso, como também fazer parte delas, exercendo as mesmas funções dos jornalistas em contexto de estágio.

Recorrendo às referidas técnicas de recolha de dados e informação, o presente capítulo procura resumir a história do jornal Expresso, especialmente da sua vertente digital, explicar como são produzidos os conteúdos do jornal, e, em especial, a rubrica dedicada ao jornalismo de dados, o 2:59.

### **3.1 Traços da história do jornal Expresso**

Desde o dia em que surgiu pela primeira vez nas bancas, a seis de janeiro de 1973, o semanário Expresso, fundado por Francisco Pinto Balsemão, não parou de desafiar e reinventar o jornalismo português. Foi o primeiro jornal a introduzir a fotolegenda (Expresso, 2018), bem como o primeiro semanário português a abraçar a Internet, juntando à versão em papel uma edição digital (Helder Bastos, 2007).

Logo a edição inaugural o jornal Expresso trouxe algo inédito: a fotolegenda em destaque na primeira página, que pretendia mostrar o mau estado das estradas nacionais no final da ditadura do Estado Novo. Onze meses depois, surge também na primeira página um mapa, que mostrava onde e como se concentravam as Forças Armadas na madrugada de 25 de abril de 1974. O referido mapa é considerado o ponto de partida para o que são hoje as infografias que integram o jornal em papel e digital (Expresso, 2018).

Mais tarde, em 1997, o Expresso cria o seu *website* e, de acordo com Helder Bastos (2007), foi o primeiro semanário em Portugal a fazê-lo. Nele começou a publicar, de forma integral e gratuita, as suas edições em papel. O início do milénio foi marcado por avanços e recuos nesta nova vertente do jornal. No final de 2000, o Grupo Impresa reforça a sua aposta *online* – isto é, nas plataformas do Expresso e da SIC –, de forma a “*tirar o máximo partido das sinergias dos investimentos do grupo no setor online*”, segundo Bastos (2007). Contudo, as receitas publicitárias adquiridas não estavam a corresponder às expectativas e, em menos de um ano, surge a necessidade de reestruturar a redação. Em termos editoriais, significou o fim da atualização de notícias *online*. E na redação, resultou no despedimento de quase duas dezenas de jornalistas (Bastos, 2007). Este cenário *online*, pouco animador, também se fez sentir nos outros *media* digitais em Portugal. Alguns chegaram mesmo a fechar portas. Com o investimento publicitário em queda e o orçamento reduzido para realizar novos projetos digitais, o Expresso foi, passo a passo, fechando os seus conteúdos, para que o público comesse a pagar por eles. Esta foi uma estratégia adotada em 2004, não só pelo Expresso como também por outros *media*, como é o caso do jornal Público.

Motivados pelo aumento do número de portugueses a navegar na Internet, em meados de 2006 os *media* portugueses voltaram a investir em novos projetos digitais. Nessa altura, o Expresso iniciou a sua aposta no *podcast*. E apenas um ano depois, anuncia renovar a sua existência digital, investindo no que traz de novo a Internet: a convergência de

diferentes meios e a interatividade. Começou a apostar fortemente em conteúdos multimédia, como em *podcasts*, fotogalerias e vídeos, bem como na interação com os leitores (Helder Bastos, 2007). Em 2011, o Expresso dá início à sua aposta em aplicações para dispositivos móveis, lançando ferramentas digitais compatíveis com os iPad e os *tablets* com sistema operativo *android* (Oliveira, 2012). Um novo e inovador projeto digital do semanário germina em maio de 2014: o Expresso Diário, primeiro jornal digital pago do país concebido para sair de segunda à sexta, às seis da tarde, com uma análise dos temas mais importantes do dia (Rodrigues, 2015). O Expresso Diário surge da necessidade de acompanhar o ritmo noticioso que um semanário não é capaz de assegurar, segundo conta José Cardoso, editor adjunto do Expresso Diário, em entrevista exploratória à investigadora. E surge também da necessidade de avançar com os conteúdos pagos (Rodrigues, 2015). Publicá-lo ao final do dia foi uma decisão tomada tendo em conta as horas vagas dos leitores, segundo José Cardoso. Um ano depois, uma nova rubrica – produzida apenas em formato de vídeo – é introduzida neste jornal diário *online*, o 2:59. Este projeto, que procura explicar o mundo a partir do cruzamento de diversos dados, é publicado uma vez por semana, primeiro nos temas principais do Diário e depois na secção, que tem o seu nome, no *site* (Expresso, 2016).

Atualmente dirigido por Pedro Santos Guerreiro, o jornal que fez nascer o Grupo Impresa – do qual também fazem parte a SIC e a Blitz – tornou-se num marco de referência do jornalismo português, defendendo que é preciso “liberdade para pensar”. De Paço de Arcos para o mundo, o Expresso conquistou uma posição de renome em Portugal: é o jornal em papel mais vendido no país e o Expresso Diário, por sua vez, é líder na circulação digital paga (APCT, 2018).



### **3.2 Do Papel ao Digital: uma breve caracterização**

O experiente semanário Expresso, que todos os sábados chega às bancas, investiu – e continua a investir – nas suas, mais recentes, plataformas digitais: o *site* e o diário. São três plataformas distintas, com traços e periodicidades próprias. O jornal em papel deixa a sua marca não só pelo formato, como também pelo conteúdo. E as características tanto do diário, como do *site* do Expresso também os diferenciam dos demais. Segue-se uma breve descrição de todas estas – e outras – particularidades destas plataformas que, hoje, juntas fazem respirar o jornal Expresso.

Atualmente, o semanário Expresso é impresso segundo o formato de *Berliner* – tem cerca de 47 centímetros por 31,5 (Figueira, 2012) – e é composto por dois cadernos. No primeiro, e principal, encontram-se notícias sobre política, assuntos internacionais, sociedade, desporto e opinião. E o segundo apresenta apenas temas sobre economia: turismo, emprego, imobiliário, bolsa, entre outros. A Revista E também integra o jornal e é na sua capa que vem o *voucher* – um pequeno código – que dá acesso às edições do Expresso Diário da respetiva semana. O referido *voucher* promove a convergência de meios, do papel e do diário digital, porque, segundo José Cardoso, a existência do Expresso Diário “*também sustenta um bocadinho as vendas do semanário*”, as quais estão a cair ano após ano (APCT, 2018) – tal como foi referido no primeiro capítulo.

No meio digital, o *website* é atualizado permanentemente, sobretudo, em sete secções: política, sociedade, economia, internacional, cultura, desporto e opinião. Neste espaço virtual, há destaque para vários projetos que nascem no seio digital do Expresso. Os *podcasts* são um deles. A “Comissão Política”, por exemplo, é um *podcast* onde é possível ouvir, todas as semanas, um debate entre jornalistas e editores sobre os principais temas de política que marcam a atualidade. E no *podcast* “A Beleza das Pequenas Coisas” é possível conhecer mais sobre as personalidades ímpares da sociedade e da cultura portuguesa e internacional. Entre os projetos digitais há também destaque para a secção multimédia, onde são publicadas reportagens compostas por diferentes meios – como texto, vídeo, áudio, ilustrações e infografias –, mas sem periodicidade definida. E surge também uma ligação para os documentários do Expresso, que possuem igualmente traços multimédia e também não têm periodicidade. A rubrica 2:59, que mostra como o jornalismo de dados pode explicar o mundo através de um vídeo, também está em grande plano. Outro projeto que se evidencia é o Expresso Curto, uma síntese da atualidade nacional e internacional redigida pelos editores do jornal, que é publicada todos os dias úteis às nove

da manhã e passível de ser subscrita pelos leitores. O guia do Expresso “Boa cama, boa mesa”, que reúne os melhores alojamentos e restaurantes de Portugal, também é um projeto promovido no *site*, assim como as iniciativas do jornal, nomeadamente a “*Global Investment Challenge*” e a “*Entrepreneur of the Year*”. As ligações aos classificados do Expresso referentes ao emprego e ao imobiliário também estão presentes no *site*.

O Expresso Diário, o jornal digital pago publicado de segunda à sexta, é dividido em sete segmentos diferentes: o dia num minuto, os temas principais, opinião, país, mundo, cultura e lazer. O primeiro segmento, o dia num minuto, reúne entre oito a 13 sínteses de notícias que marcam o dia. Geralmente são desenvolvidos entre seis a oito temas principais que marcam a atualidade, os quais fazem parte da capa do jornal digital. O terceiro segmento conta com as opiniões, sobre temas atuais, de três ou mais autores: os que escrevem com regularidade para o jornal, como Henrique Raposo e Daniel Oliveira, e os editores do próprio jornal. As notícias importantes a nível nacional e internacional – geralmente já publicadas no *site* – também são atualizadas para os segmentos o País e o Mundo, respetivamente. E para as secções de Cultura e Lazer, regra geral, é desenvolvido apenas um tema.

A evolução do jornal Expresso é notória. Durante mais de 20 anos foi um semanário apenas. Depois, procurou adaptar-se ao meio digital, investindo calculosamente em vários projetos no seu *website*. E, recentemente, lançou um projeto inédito em Portugal, um jornal digital pago bem definido, o Expresso Diário. Estas três plataformas têm características distintas, mas partilham o mesmo rigor e qualidade jornalística. E refletem a capacidade que o Expresso tem de se adaptar ao universo digital.

### **3.2.1 Estratégias de negócio do Expresso Digital**

Atualmente, o Expresso aposta, essencialmente, em duas estratégias de negócio distintas para as suas plataformas digitais. No *site*, os conteúdos são gratuitos. Já no Expresso Diário, o jornal digital, geralmente há um custo associado à leitura integral das notícias.

No *site* do Expresso, a estratégia de negócio apresentada é uma combinação de *freemium*, isto é, os conteúdos publicados podem ser lidos sem qualquer custo, com receitas publicitárias, que são tanto maiores quanto maior for o número de visitantes do *site*. A referida publicidade geralmente surge ao abrir as notícias e nas margens laterais das mesmas. Para além do referido, por vezes, também são desenvolvidos artigos patrocinados, que geralmente exploram as utilidades de um novo produto ou serviço de uma determinada marca. Através do *site* também é possível aceder à loja virtual do Expresso, que é um exemplo de uma aposta *e-commerce* do Grupo Impresa. Neste espaço digital, não só surgem as várias ofertas de assinatura digital do Expresso como também filmes, séries, livros (alguns exclusivos), o guia “Boa cama, boa mesa”, entre outros.

A estratégia de negócio do Expresso Diário também é combinada. Para além da publicidade que pode surgir na abertura do jornal e no rodapé dos temas, este jornal digital é pago. Isto quer dizer que para ter acesso à leitura integral do Expresso Diário de segunda à sexta, o público tem de realizar um investimento monetário. Ou adquire o semanário ao sábado e através do *voucher* – presente na capa da Revista E – acede a todas as edições do Expresso Diário da semana seguinte. Ou seleciona uma das opções de subscrição do Expresso Digital, que dão acesso à versão digital do semanário (que inclui os dois cadernos e a Revista E) e também a todas as edições do Expresso Diário de segunda a sexta. Caso não realize nenhuma das opções mencionadas, o público apenas tem acesso à parte inicial dos temas publicados no Expresso Diário, os quais são promovidos no *site*.

Estas são apenas as principais estratégias de negócio associadas às plataformas digitais do Expresso, que promovem – ou tentam promover – a sua sustentabilidade financeira. Há outras, mais complexas, que a investigadora não tem informação suficiente para as descrever, como é o caso das parcerias entre o Expresso e o BPI na vertente imobiliária do jornal.

### **3.3 As rotinas produtivas: da seleção de temas à publicação**

Todo o processo de produção de notícias – modulado no Apêndice 2, segundo a notação BPMN (*Business Process Model and Notation*) e com recurso à ferramenta *Visual Paradigm 15.1* – tem um princípio, que se inicia quando há temas com interesse público e novidade tal que merecem ser desenvolvidos, e um fim, que corresponde à fase de publicação dos mesmos. E tem também atores que intervêm diretamente no processo: os editores, que decidem quando e onde publicar os temas, os jornalistas, que os desenvolvem, os fotojornalistas, que os complementam com fotografias, entre outros.

Todas as semanas, a produção do jornal Expresso – tanto do semanário, como do diário e do *site* – afigura-se num verdadeiro desafio. Tudo começa com uma ideia, um tema com novidade passível de ser desenvolvido, que pode ter diferentes origens. Podem ser propostos pelos jornalistas, podem surgir nas agendas das agências noticiosas, principalmente da Lusa, e também da leitura de jornais diários, segundo refere José Cardoso. As dificuldades começam a surgir na hora de distribuir os temas pelas várias plataformas. A fraca durabilidade das notícias ditas “em primeira mão” é, segundo José Cardoso, o problema central. Com um fluxo noticioso permanente no meio digital, torna-se “ *muito difícil para um semanário ter notícias em primeira mão*”, conta o editor. E desenvolver o número de temas adequado para preencher o jornal em papel é também uma dificuldade.

Face a esta instabilidade, a forma como os temas são geridos no seio digital também muda de plataforma para plataforma. “*O site está sempre a ser alimentado em contínuo*”, refere o editor, e a seleção e gestão de temas passa, sobretudo, pelas mãos do coordenador multimédia Paulo Luís de Castro, mas também pelos jornalistas da secção. O Expresso Diário para além de ter um objetivo próprio – o de aprofundar os temas do dia no próprio dia – possui um limite de temas principais (entre seis a oito), o que torna a sua gestão num processo “ *muito plástico [flexível]*”, conta José Cardoso, editor adjunto deste jornal, que partilha a sua gestão com o editor executivo Germano Oliveira. Tudo pode mudar de uma hora para a outra. Os temas podem ser decididos de manhã, de véspera, ou até alguns dias antes, mas nada garante que os sete temas definidos com antecedência coincidam com os sete que são publicados no Expresso Diário ao final da tarde. E há sobretudo duas razões que explicam este cenário volátil, explica o editor. Por vezes surgem, ao longo do dia, histórias importantes que acabam por substituir as definidas inicialmente. E muitas vezes as notícias escolhidas são “queimadas por outros *sites*”, levando os jornalistas a publicá-

las no *site* do Expresso. Para fazer face a esta instabilidade do jornal digital, José Cardoso conta que é habitual escrever primeiro uma breve notícia para o *site* sobre determinado tema e depois “*desenvolvê-la de outra maneira, com alguma mais valia, para o diário*”.

A distribuição de temas pelos jornalistas nem sempre é igual, mas, segundo José Cardoso, há um princípio a seguir. Os temas seleccionados para o Expresso Diário são, num primeiro momento, atribuídos aos jornalistas que os seguem, isto é, que os acompanham de perto e que, por isso, estão mais habilitados para escrever sobre os mesmos. Quando não é possível, os temas são entregues aos jornalistas que, naquele dia, têm mais disponibilidade para os desenvolver. E a mesma lógica predomina no *site*. Em termos gerais, as notícias são desenvolvidas por 56 jornalistas, que tanto escrevem para o semanário como também para a vertente digital do jornal – *website* ou diário –, mas com frequências diferentes.

Depois de atribuídos os temas, os jornalistas possuem liberdade para os desenvolver da forma que considerarem mais adequada. Este processo, o de construção da história, também não é linear, mas pode-se resumir nas seguintes fases, que podem assumir esta ou outra sequência: os jornalistas pesquisam informação em fontes oficiais, sobretudo, na Internet; constroem guiões informativos; contactam as fontes; seleccionam as informações mais relevantes tanto da sua própria pesquisa como das entrevistas realizadas; redigem uma história; solicitam fotografias, gráficos e/ou infografias aos respetivos profissionais; e inserem a história e os seus conteúdos no sistema de informação adequado: no *Content Station*, se a história sair no semanário ou no diário, e no *Backoffice* se for publicada no *site*. No final do processo estão novamente os editores, os quais leem e analisam a história, corrigem-na quando necessário e publicam-na nas respetivas plataformas.

Como o processo de desenvolvimento de notícias nem sempre é igual, há nuances a serem consideradas. Na fase de pesquisa de informação, os jornalistas podem, por exemplo, consultar o sistema de arquivo do jornal, o *Elvis* – gerido pela Gesco, uma outra marca do Grupo Impresa – e pesquisar aí artigos que já foram publicados sobre o tema que estão a desenvolver. Quando as histórias são baseadas em dados, há que introduzir a fase de tratamento, análise e interpretação de dados, bem como as ferramentas de trabalho que suportam o processo, como o *Excel*. Quando os jornalistas saem em reportagem para cobrir determinado acontecimento e não chegam à redação a tempo de assegurar todo o processo, este é compartilhado com os editores ou outros jornalistas.

Após a publicação das notícias nas plataformas digitais, surge uma outra fase, a de promoção das mesmas nas redes sociais, como no *Facebook*, no *Twitter*, no *Linkedin* e no *Instagram*. Este trabalho é realizado por uma equipa que se dedica essencialmente à gestão destas redes, não só do Expresso, mas também da SIC Notícias e da revista Blitz. Para estes profissionais, a ferramenta *Google Analytics* é bastante útil, por várias razões. É através dela que é possível conhecer, em tempo real, quantos leitores estão a navegar nas plataformas e quantos as acedem pelas diferentes vias: pelas redes sociais e diretamente pelo *site*. E permite descobrir também quais são os temas mais e menos lidos, entre outras informações. Para o editor José Cardoso, há duas formas de ler estes dados. Por um lado, pode-se voltar a promover nas redes sociais os temas mais lidos. E, por outro, pode-se também republicar ou destacar os temas importantes que não estão a ser lidos. Estas informações provenientes do *Google Analytics* podem ter outras finalidades, nomeadamente na tomada de decisão. Ao descobrir as tendências de leitura contidas nos dados, os editores procuram novas histórias sobre esses temas, que têm especial interesse para o público. “*Notícias que tenham a ver com imobiliário, com consumo, telecomunicações, energia e conselhos de poupanças são muito lidas. E, portanto, tentamos ter muitas e boas notícias nessas áreas*”, explica o editor. Mas mesmo com estas pistas, a produção de notícias não é fácil, porque, segundo o editor, “*já praticamente tudo foi criado*”. O grande desafio hoje é encontrar novas perspetivas sobre estes temas.

Para contornar estas dificuldades geradas pela instabilidade do noticiário, o Expresso constrói, todos os dias, uma relação simbiótica entre as diferentes plataformas. O digital suplementa o papel e o papel, por sua vez, também dá vida ao digital. Isto quer dizer que é no digital – no *site* e no diário – que são publicados os temas quando não deixam de ser exclusivos até sábado. E também quer dizer que os temas produzidos para o papel são publicados no seio digital – nomeadamente no *site* – aos fins de semana, por exemplo, quando há menos jornalistas na redação.

### **3.3.1 Como é produzida a rubrica 2:59?**

Em meados de 2015 surge o 2:59, uma nova rubrica do Expresso Digital especialmente dedicada ao jornalismo de dados. Antes de descrever o processo inerente à construção da rubrica – que vai ao encontro do descrito na literatura embora com algumas nuances – há que regressar ao passado e perceber como tudo começou, como esta ideia surgiu.

Foi na mente de Pedro Santos Guerreiro, o atual diretor do jornal Expresso, que se formou a ideia inicial, a qual acabou por dar nome à rubrica: o número 2:59. Este número quebrado – que é uma estratégia utilizada pelas grandes superfícies para estipular preços – pareceu a Pedro Santos Guerreiro o ideal para entrar na mente do público como uma marca, conta o diretor em entrevista informal à investigadora. Sabia que queria inovar, fazer nascer um novo projeto no Expresso, mas ainda não sabia o que este número iria significar. Ao investigar as estatísticas sobre o tempo médio de atenção do público ao visualizar conteúdos audiovisuais – que é de apenas um minuto – percebeu que o tal número, o 2:59, poderia significar isso mesmo: dois minutos e 59 segundos. Mas, nessa altura, faltava ainda saber qual o conteúdo a inserir na rubrica, conta o diretor.

A ideia inicial começou a ganhar forma e conteúdo quando Pedro Santos Guerreiro se deixou inspirar por dois vídeos. Um deles mostrava Barack Obama, o ex-presidente dos Estados Unidos da América, a dar uma entrevista envolvido num fundo preto, no qual iam aparecendo infografias relacionadas com os temas que abordava. E no outro vídeo, o estatístico Stephen Stigler explicava um tema com base em dados, estatísticas próprias, e onde os gráficos e animações acompanhavam os seus movimentos. Foi nestas duas fontes de inspiração que Pedro Santos Guerreiro encontrou o conteúdo que procurava. Assim, definiu que num fundo preto, um jornalista iria contar uma história baseada em dados e que, no desenrolar da mesma, iam surgindo infografias que podiam acompanhar os movimentos do jornalista.

No 2:59 que explica como se faz um 2:59, publicado em julho de 2016, Pedro Santos Guerreiro faz uma breve descrição do processo por detrás da rubrica:

*“Nós pegamos num conjunto de dados e transformamo-los numa história. É assim que tudo começa. Corrigimos os dados, escrevemos um guião e depois descemos até ao estúdio [da SIC] e com um fundo negro e, sem teleprompter, temos que dizer um texto que decorámos em menos de dois minutos e 59 segundos. O que nem*

*sempre é fácil. E só depois disso é que entra todo o grafismo, para que tudo funcione”* (Expresso, 2016).

Neste pequeno excerto do guião deste 2:59, o atual diretor do jornal Expresso refere as principais fases de produção da rubrica – as quais estão moduladas no Apêndice 3. Tudo começa quando há um tema de interesse público, sobre o qual existem dados credíveis passíveis de serem recolhidos. Ou quando há um conjunto de dados recentes que podem ser explorados. Estes dois pontos de partida para dar início ao processo de jornalismo de dados correspondem aos que Bradshaw (2011b) refere na literatura. Os temas que surgem são apresentados à coordenadora multimédia Joana Beleza, que decide se são ou não adequados à rubrica. Depois, o processo de jornalismo de dados segue com naturalidade: os dados são recolhidos, limpos, analisados e interpretados pelos jornalistas. Quando surge a fase de redação do guião, os jornalistas devem ter em conta que o estilo do discurso deve ser adequado ao meio audiovisual e ser, portanto, mais coloquial e fluído. E esta particularidade do 2:59, o seu formato audiovisual, é o que molda a forma como a história é contada, moldando, portanto, a fase de comunicação inerente ao processo de jornalismo de dados descrito por Bradshaw (2011a). Este ponto, em específico, revela a faceta multimédia do Expresso e, portanto, a importância de ser versátil para trabalhar em diferentes meios. Depois de redigir o guião e de ajustar os dados numa folha de *Excel*, os jornalistas enviam o material à coordenadora Joana Beleza que revê e faz os ajustes que considerar importantes. Antes do dia da gravação, a coordenadora recorda aos jornalistas as técnicas não verbais que são adequadas e as que se devem evitar. Além do referido, coordenadora também é responsável por agendar as gravações do 2:59 nos estúdios da SIC. Esta particularidade, em especial, revela a convergência mediática existente no Grupo Impresa, que se traduz numa parceira entre o Expresso e a SIC, a qual autoriza a utilização dos recursos audiovisuais técnicos e humanos da SIC para gravar a rubrica do jornal. No dia da gravação, os jornalistas são, geralmente, acompanhados pelo realizador André Atayde, e, já no estúdio, deparam-se com um fundo verde e não negro. E hoje já têm o apoio do teleponto que vai exibindo o guião à medida que é dito. Seguindo as indicações do realizador, os jornalistas podem gravar uma, duas, três vezes, as necessárias, até que surja uma versão sem gafes e dita no tempo certo. As gravações do 2:59 são depois enviadas ao *designer* gráfico Carlos Paes que desenha e introduz no vídeo todo o grafismo. Por fim, a última versão da rubrica só é publicada no Expresso Diário quando os seus editores o decidirem.





## **4 Seis meses no seio do jornal Expresso**

### **4.1 Objetivos do estágio**

Os objetivos de aprendizagem e investigação propostos no âmbito do estágio na redação do jornal Expresso passam por:

- Estudar o estado-da-arte do jornalismo de dados;
- Melhorar a destreza na utilização de *software* de tratamento e visualização de dados;
- Estudar, modular e analisar os processos produtivos na redação do Expresso Digital;
- Identificar as principais fontes de informação e base de dados acedidas;
- Investigar a viabilidade de construir uma base de dados institucional no jornal Expresso;
- Estudar a estrutura de *E-business* e *E-commerce* inerente ao jornal digital;
- Investigar e refletir sobre os principais sistemas de informação utilizados;
- Estudar e modular os processos de construção das peças multimédia na rubrica 2:59;
- Otimizar os processos produtivos na redação do Expresso Digital;
- Investigar a perceção dos leitores relativamente ao jornalismo de dados através de questionários;
- Caso de estudo do Expresso Digital – Investir e Inovar em Jornalismo de Dados:
  - Investigar as mudanças no contexto organizacional quanto à introdução do jornalismo de dados;
  - Analisar as competências e perfis dos profissionais que se dedicam ao jornalismo de dados;
  - Investigar a importância dos produtos resultantes do jornalismo de dados para os leitores;

Tendo em conta a relevância, dimensão e alinhamento de alguns objetivos, optei por desenvolvê-los em outros capítulos e subcapítulos. A correspondência entre objetivos e capítulos de desenvolvimento está descrita na tabela 4.1, que se segue. Os não mencionados serão abordados no diário de bordo.

**Tabela 4.1** – Correspondência entre objetivos e capítulos de desenvolvimento

**Fonte:** Elaboração própria

<b>Objetivos</b>	<b>Capítulos de desenvolvimento</b>
Estudar o estado-da-arte do jornalismo de dados	Capítulos 1 e 2
Estudar, modular e analisar os processos produtivos na redação do Expresso Digital	Subcapítulo 3.3 Modulação – Apêndice 2
Estudar a estrutura de E-business e E-commerce inerente ao jornal digital	Subcapítulo 3.2.1
Estudar e modular os processos de construção das peças multimédia na rubrica 2:59	Subcapítulo 3.3.1 Modulação – Apêndice 3
Investigar a perceção dos leitores relativamente ao jornalismo de dados através de questionários	Subcapítulo 5.2.1 Apêndice 5
Caso de estudo do Expresso Digital – Investir e Inovar em Jornalismo de Dados	Capítulo 5

## **4.2 Diário de Bordo**

O diário de bordo, que se segue, reúne os principais desafios, aprendizagens e projetos que desenvolvi no decorrer do estágio curricular na redação do Expresso Digital – situada no Edifício do Grupo Impresa em Paço de Arcos –, que teve início a 18 de setembro de 2017 e término a 18 de março de 2018. Na avaliação final do estágio, realizada pelo meu supervisor Germano Oliveira, obtive a classificação de 18,2 valores.

### **4.2.1 Acolhimento**

Abraçar o estágio no jornal Expresso pode resumir-se numa só palavra: mudança. Mesmo antes de o iniciar, tudo mudou à minha volta. Dia 16 de setembro de 2017, mudei de morada e acomodei-me na acolhedora e, até então, desconhecida vila de Oeiras. Instalei-me numa nova casa e reaprendi a dividi-la com três pessoas da minha geração. Nunca tinha vivido tão perto do mar, nem tão perto da agitada capital, Lisboa. Fiquei só, entregue a mim mesma. Logo fui conhecendo os cantos à vila e, em especial, o caminho para aquele que seria o meu local de trabalho, o Edifício do Grupo Impresa, em Paço de Arcos. Levei cerca de 25 minutos a percorrê-lo pelo meu próprio pé e decidi que era assim que iria fazê-lo nos seis meses seguintes.

Dois dias depois da minha chegada à nova morada, iniciei o meu percurso enquanto estagiária do Expresso Digital. O dia 18 de setembro de 2017 ficará para sempre marcado como o meu primeiro dia de estágio curricular do Mestrado de Sistemas de Informação de Gestão, o dia em que iniciei o meu caminho profissional enquanto futura jornalista. Cheguei por volta das 9 horas da manhã ao Edifício do Grupo Impresa e dirigi-me à receção, onde me foi atribuído um cartão de estagiária. De elevador subi até ao quarto andar e assim que entrei na redação do Expresso procurei pela secretária do diretor Ana Bela Vieira, para que me pudesse encaminhar para a secção digital. O espaço é amplo de tal modo que, numa primeira visita, não é possível identificar as secções do jornal. Assim que cheguei ao destino, conheci o coordenador multimédia, Paulo Luís de Castro, bem como os jornalistas do digital, que estavam a trabalhar naquele dia. E instalei-me, junto deles, numa secretária, naquele que seria o meu posto de trabalho nos seis meses seguintes. Nesse mesmo dia, foi-me atribuído, pela secretária do diretor, um *email* pessoal e as credenciais necessárias para aceder aos principais sistemas de informação que suportam os processos produtivos da redação do Expresso, o *Content Station* e o *Backoffice*. Logo fui desafiada pelo coordenador multimédia a escrever uma breve notícia

para o *website* sobre a visita do primeiro ministro da Roménia, Mihai Tudose, a uma das cidades mais devastadas pela tempestade que atingiu o país nessa época, Timisoara. Esta notícia – como todas as que desenvolvi ao longo do estágio – não foram assinadas com o meu nome, mas com a assinatura do Expresso, porque somente os jornalistas que possuem carteira profissional o podem fazer.

#### **4.2.2 Projetos Desenvolvidos**

##### ***Dia num Minuto do Expresso Diário***

Na primeira semana, foi-me atribuída uma tarefa diária pelo editor adjunto José Cardoso, que abracei até janeiro de 2018: produzir o “Dia num minuto” do Expresso Diário. Como o nome indica, esta secção propõe-se a sintetizar as notícias importantes do dia – entre 8 a 13 – de forma a que o público as possa ler em apenas um minuto. O primeiro plano desta secção é ocupado por uma frase com destaque no dia e, em seguida, surgem uma ou duas notícias que evidenciam números. Geralmente, à medida que seleccionava as notícias ao longo do dia – publicadas no *site* do Expresso –, sintetizava-as em 4 ou 5 linhas. E quando grande parte das notícias já estavam resumidas, solicitava à secção de fotografia do Expresso – assegurada pela fotojornalista Ana Baião e pelo coordenador de fotografia João Carlos Santos, entre outros profissionais – uma fotografia para a frase do dia e duas para as notícias que, segundo a minha perspetiva, mereciam especial destaque. Para contactar estes e outros profissionais usei com frequência o sistema de comunicação interno existente na redação, o denominado VOIP (*Voice Over Internet Protocol*), que permite enviar mensagens e telefonar aos colegas da redação. É de salientar que as notícias eram reformuladas sempre que surgisse uma atualização. O desafio era terminar o processo antes das seis da tarde, para que os editores do jornal pudessem fazer uma breve revisão antes de publicar o Expresso Diário.

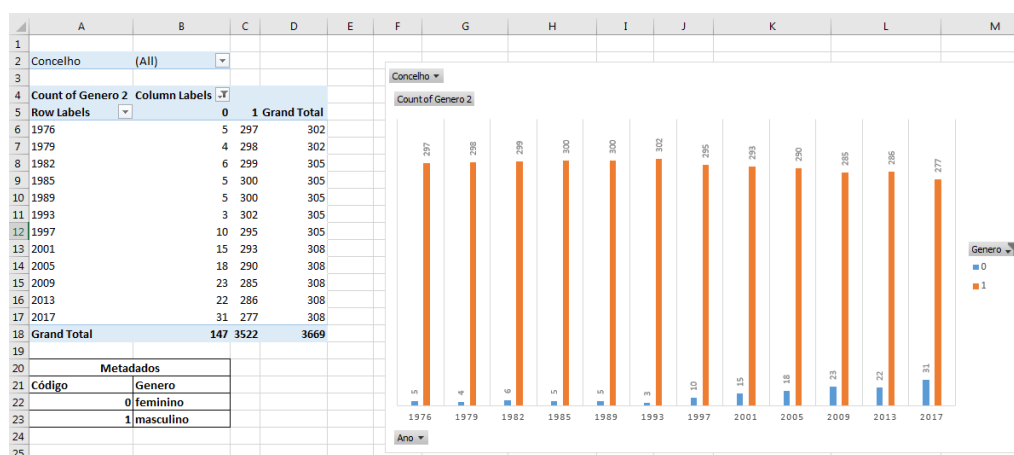
##### ***O primeiro trabalho de jornalismo de dados: as mulheres nas autarquias***

O meu futuro na redação do Expresso começou a ganhar forma quando Germano Oliveira, o meu supervisor e editor executivo do Expresso, me apresentou à jornalista Raquel Albuquerque, que se dedica ao jornalismo de dados, e que, por ele, foi nomeada mentora do meu estágio. Para que a jornalista pudesse acompanhar o meu trabalho, passei a fazer um horário semelhante ao seu: das 10 às 18 horas – até então cumpria um horário ligeiramente diferente, das 9 às 17 horas, mas igualmente flexível.

A jornalista, que já esperava a minha chegada na redação, logo me desafiou a encontrar estratégias para analisar um conjunto de dados já recolhidos por si junto da Secretaria Geral do Ministério da Administração Interna. A folha de cálculo do *Excel* (em formato .xlsx) que me foi enviada pela jornalista – via *email* – compilava informações sobre todos os presidentes eleitos nas 12 eleições autárquicas que decorreram até hoje em Portugal. O foco de análise proposto por Raquel Albuquerque era claro: as mulheres eleitas presidentes nas autarquias.

Perante um ficheiro de *Excel* com 12 folhas – uma para cada ano de eleições autárquicas – comecei por compilar numa folha só todos os dados sobre os autarcas e ex-autarcas: os anos de eleição, os municípios, os nomes completos e os respetivos partidos. Em seguida, procedi à limpeza dos nomes – que nem sempre eram referidos do mesmo modo – de forma a homogeneizá-los. Para conseguir compreender quantas mulheres foram eleitas ao longo do tempo, inseri filtros nas colunas existentes e adicionei mais uma, a qual referenciei de “Género”. Codificar os géneros – 0 para as mulheres e 1 para os homens – foi a estratégia que encontrei para os diferenciar, a qual foi aprovada pela jornalista. Estes foram os primeiros passos que percorri para tratar a base de dados, os quais coincidem com as primeiras fases do processo de jornalismo de dados referenciadas na literatura: nomeadamente as fases de recolha (a partir de diversas fontes de dados), limpeza e compilação de dados referidas por Bradshaw (2011b).

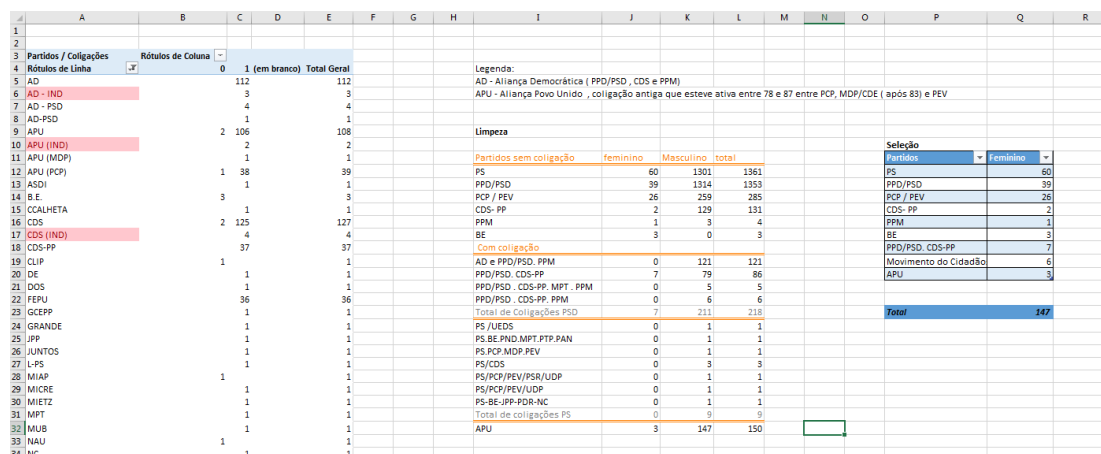
O passo seguinte diz respeito à análise dos dados, a qual foi realizada segundo várias perspetivas. Estudei a evolução do número de mulheres eleitas presidentes das autarquias ao longo de 41 anos de democracia. Para realizar a referida análise, isolei numa folha de *Excel* as variáveis referentes ao ano de eleição e ao género, anteriormente codificado. Na tabela dinâmica representada na figura 4.1, foi possível desagregar a variável género em duas colunas – em feminino (0) e masculino (1) – e, assim, perceber quantas mulheres e quantos homens foram eleitos nas 12 eleições autárquicas ocorridas até então. O gráfico de barras, também dinâmico, permitiu mostrar algo que os números, por si só, já diziam: houve – e há – muito mais homens do que mulheres eleitas presidentes nos 308 municípios portugueses. Ficaram evidentes as brutais diferenças de género na ocupação destes cargos políticos, apesar do ligeiro aumento do número de mulheres a ocupá-los nas duas últimas décadas. E estas foram as primeiras informações que extrai da análise dos dados.



**Figura 4.1** – Tabela e gráfico dinâmicos que mostram a evolução do número mulheres e homens eleitos presidentes desde 1976 até 2017

**Fonte:** Elaboração própria (folha de cálculo do *Excel*, em formato .xlsx)

Outra perspetiva explorada nos dados diz respeito ao total de mandatos femininos correspondentes a cada cor política – PS, PSD, PCP, BE, entre outros –, desde 1976 até à atualidade, como mostra a figura 4.2.



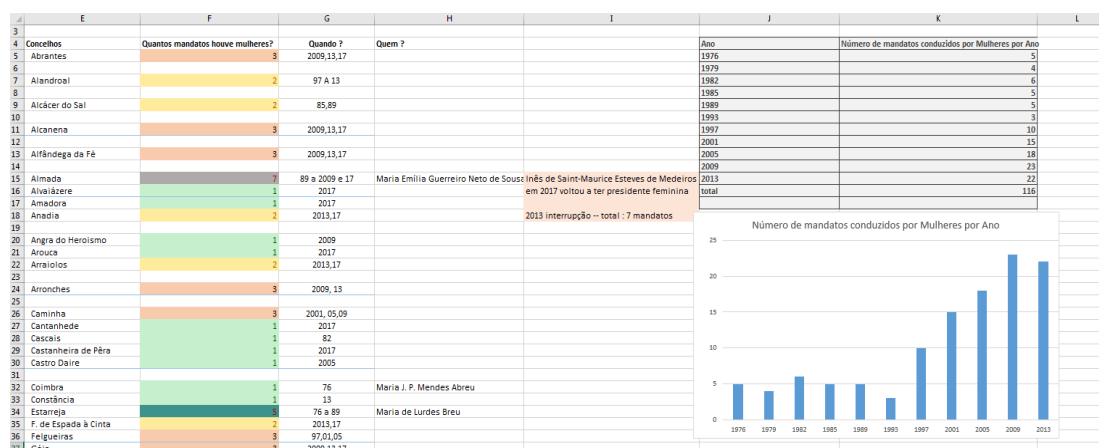
**Figura 4.2** – Número total mandatos (femininos e masculinos) por partidos e coligações de autarcas (diferenciado por género)

**Fonte:** Elaboração própria (folha de cálculo do *Excel*, em formato .xlsx)

Depois de filtrar as variáveis correspondentes aos partidos políticos e ao género, apercebi-me da quantidade imensa de partidos e coligações que já existiram do país. E rapidamente esta constatação revelou-se uma dificuldade. Os nomes dos partidos políticos e das coligações não estavam todos escritos do mesmo modo, o que evidenciou uma necessidade de os uniformizar manualmente. Face à quantidade de dados, decidi selecionar apenas os partidos (e coligações) que obtiveram um maior número de mandatos no global. E, posteriormente, ordenei de forma decrescente o número de mandatos femininos, para perceber qual é o partido que, até hoje, possuiu mais mulheres a presidir

autarquias. Neste caso, o Partido Socialista lidera a lista, somando 60 mandatos femininos.

A análise com maior destaque – que acabou por definir o fio condutor do trabalho – foca os concelhos que já tiveram pelo menos uma mulher sentada na cadeira da Presidência. Para realizar esta análise, comecei por isolar as seguintes variáveis: “concelho”, “ano de eleição”, “nome próprio” e “género”. E, em seguida, filtrei todas as mulheres, para perceber quais eram, numa primeira instância, os concelhos que tiveram pelo menos uma mulher na liderança. Reduzida a lista a apenas 62 concelhos liderados por mulheres, prossegui para a análise manual, na qual contabilizei quantos mandatos femininos cada concelho obteve. E também filtrei os nomes das autarcas e ex-autarcas, bem como os respetivos anos de eleição – tal como evidencia a figura 4.3. Esta análise mostrou uma tendência: foram várias as mulheres que assumiram as rédeas das autarquias anos a fio, 5 e 6 mandatos consecutivos.



**Figura 4.3** – Número de mandatos femininos por concelho (à esquerda) e evolução do número de mandatos conduzidos por mulheres desde 1976 até 2017 (à direita)

**Fonte:** Elaboração própria (folha de cálculo do *Excel*, em formato .xlsx)

Partindo desta pista, procurei contactar estas mulheres. Consegui entrevistar três ex-autarcas, via telefone e via *email*: Francelina Chambel, uma das cinco primeiras autarcas eleitas em 1976, que presidiu a câmara do Sardoal até 1993 (5 mandatos consecutivos); Maria de Lurdes Breu – também uma das cinco primeiras – que liderou o município de Estarreja até 1993, perfazendo um total de 5 mandatos; e Maria Irene Barata, que presidiu o concelho de Vila de Rei entre 1989 e 2013 (6 mandatos consecutivos). Conhecer a história destas mulheres foi uma experiência acutilante. Conhecer de perto suas emoções, os seus receios, as suas conquistas foi, sem dúvida, uma grande honra. Não bastou um *email* ou uma chamada telefónica para conhecer todos os aspetos importantes na vida



política destas mulheres. Foram vários os contactos trocados. Poder inserir as suas histórias naquele que seria o meu primeiro grande trabalho de jornalismo de dados foi um verdadeiro desafio.

Depois de recolher os testemunhos junto das ex-autarcas e de analisar e interpretar os dados, prossegui para a fase seguinte: a redação da história. Selecionar tanto os dados mais relevantes como as informações mais pertinentes das entrevistas foi um primeiro passo para a construção da história. E, depois, a interpretação dos dados e os testemunhos foram-se entrelaçando, sob orientação da jornalista Raquel Albuquerque. O seu início focou as mulheres que mais mandatos consecutivos conquistaram – como foi referido anteriormente. Com o intuito de inserir três gráficos interativos na história, enviei à secção de infografia do Expresso – em especial à infografista Sofia Miguel Rosa – os dados referentes à evolução dos mandatos conquistados por mulheres e por homens ao longo de 12 eleições autárquicas. Enviei também os dados alusivos aos concelhos já presididos por mulheres (incluindo o número de mandatos femininos de cada um, bem como os anos) e relativos ao número de mandatos liderados por mulheres que cada partido político conquistou desde 1976 até à atualidade. E solicitei também fotografias à respetiva secção do Expresso para mostrar aos leitores quem são estas mulheres. Como não foi possível encontrar fotografias suficientes no arquivo do jornal, gerido pela Gesco, tive de solicitá-las às ex-autarcas entrevistadas, as quais me forneceram o material amavelmente.

Foi no dia 17 de outubro de 2017 que inseri todo o material – texto, documentos, fotografias e os três gráficos interativos (um mapa e dois gráficos de barras) – no *Content Station*, na secção dos temas principais do Expresso Diário. Depois de revisto pela editora de política do Expresso, Helena Pereira, foi, nesse mesmo dia, publicado na íntegra no Expresso Diário, parcialmente no *site* – onde foi inserido na secção dedicada às Eleições Autárquicas de 2017 (figura 4.4) – e foi também promovido nas redes sociais. No *Facebook*, em particular, a história não obteve um número de reações surpreendente: somou 14 gostos, cinco comentários e seis partilhas. Já os dados registados no *Google Analytics* mostram outro desfecho. Mostram que esta história – intitulada “Mulheres a mandar nas autarquias, procuram-se. Em 1976 eram cinco, hoje são 31” – foi visualizada por mais de duas mil pessoas no dia da publicação. E mostram também que até ao final do ano 2017 quase quatro mil e 500 leitores acederam ao artigo no Expresso Diário.



**Figura 4.4** – Interface do *site* do Expresso da história sobre as mulheres autarcas (a partir da qual é possível aceder à história integral publicada no Expresso Diário)

**Fonte:** Cópia integral de Expresso & Sofia (2017)

#### *Notícias baseadas em dados desenvolvidas para o Expresso Diário*

Muitos dos temas que redigi para o Expresso Diário foram baseados em dados. Quando o editor Germano Oliveira ou o editor ajunto José Cardoso reconheciam interesse público em dados publicados no próprio dia, desafiavam-me a analisá-los. Alguns destes temas eram sobre economia: sobre a realidade dos trabalhadores por conta própria, ou sobre o estado do emprego em Portugal, por exemplo. Outros enquadravam-se na secção de sociedade, como os motivos que estão por detrás da melhoria do bem-estar dos portugueses.

Desenvolver estes temas sob pressão, em apenas três ou quatro horas foi uma verdadeira aventura. Significou acelerar fortemente o processo produtivo. Neste aspeto, a colaboração dos editores e dos jornalistas foi crucial. Para encontrar, rapidamente, informações relevantes entre os dados, os editores verificavam e validavam as minhas análises. E para depressa conseguir entrevistar especialistas, os contactos disponibilizados pelos jornalistas foram fundamentais. Quando os prazos são curtos, torna-se difícil ter a atenção necessária para evitar erros. Torna-se difícil gerir o tempo, para que seja possível retificar toda a história, todos os cálculos efetuados, toda a ortografia. Construir histórias baseadas em dados de raiz foi, sem dúvida, um dos maiores desafios que enfrentei enquanto estagiária do jornal Expresso. E foi também o que me fez ter uma noção real do trabalho diário de um jornalista.

Para lidar com estes desafios diários, ter a humildade para pedir auxílio a outros profissionais, quando percebemos que há uma falha de conhecimento que não nos permite avançar, é importante. E para analisar dados de forma eficaz é crucial dominar as funcionalidades do *Excel* – a folha de cálculo disponível na redação. Ter um conhecimento geral de estatística foi também importante para adequar as operações à natureza dos dados e executá-las com rapidez. Em suma, esta experiência mostrou-me que, para fazer jornalismo de dados, é crucial possuir um leque de conhecimentos que abracem, sobretudo, a matemática e os sistemas de informação. E são estes conhecimentos que, tal como foi referido na literatura, caracterizam o perfil que procurei desenvolver na redação do Expresso: o de jornalista de dados.

### ***Série “Em pequeno número”***

Colaborar na segunda edição da série “Em pequeno número” foi outro grande desafio que a jornalista Raquel Albuquerque me colocou. Esta série do Expresso, publicada nas duas últimas semanas de 2017, visa contar a história de dez portugueses que atualmente vivem em países improváveis, isto é, onde quase não os há. Identificar os países onde vivem poucos portugueses – menos de 50 – no *website* oficial da Direção Geral de Assuntos Consulares e das Comunidades Portuguesas foi o primeiro passo. Encontrar portugueses que neles habitam foi o segundo e o mais difícil. Depois de vários dias de pesquisa na *web* e após diversas tentativas de contacto com portugueses através das redes sociais – *Facebook* e o *Expatblog* –, consegui entrevistar apenas três portugueses que atualmente vivem em destinos improváveis: na Guiné Conacri, no Gana e na Bolívia. Foi incrível perceber que, através da Internet, é possível reunir testemunhos de portugueses que estão a milhares de quilómetros de distância em apenas alguns dias. Escrever sobre os motivos que os levaram a emigrar, sobre suas emoções à chegada da nova morada, as suas rotinas, o dia a dia da população nativa dos respetivos países foi uma autêntica viagem, que moldou a forma como hoje escrevo histórias. A empatia é o elemento central. E foi neste projeto que aprendi a criá-la.

### ***Um dia na Assembleia da República e longe da redação***

Em janeiro de 2018 surgiu uma nova aventura. A jornalista Carolina Reis convidou-me para a acompanhar até à Assembleia da República com o intuito de cobrir o debate em torno das propostas de lei do Bloco de Esquerda e do PAN sobre a legalização da canábis para fins medicinais. Sentar-me junto de outros jornalistas, na seção destinada à

comunicação social, foi uma honra. Contar ao mundo o que ali se debateu e decidiu, o que assisti em primeira mão, foi um privilégio para mim enquanto estagiária. Ali conheci o verdadeiro sabor de ser jornalista, de ser um testemunho privilegiado dos acontecimentos. Mas também conheci a responsabilidade de recolher toda a informação necessária no momento certo para informar o público da forma mais clara e objetiva possível. Este dia, em particular, marcou verdadeiramente o meu estágio, pois foi um dos únicos em que sai da redação para estar em contacto com o real, com o acontecimento. A esmagadora maioria dos meus dias foram passados na redação, onde pesquisava informação, sobretudo na Internet, e contactava os especialistas via telefone, via *email*, ou através das redes sociais. Este cenário ilustra – ou pode ilustrar – como a Internet e as tecnologias mudaram profundamente as rotinas dos jornalistas no seio de uma redação. Hoje os jornalistas passam mais tempo junto dos computadores e da Internet, do que em reportagem.

### ***Integrar a equipa das redes sociais do Grupo Impresa***

No início de 2018, surgiu um novo desafio, desta vez, proposto pelo coordenador multimédia Pedro Monteiro: integrar a equipa que gere as redes sociais do Grupo Impresa. Na época, para além da página do jornal Expresso, inseriam-se também as páginas da SIC Notícias, da revista Visão, da Blitz e da Exame Informática. Após a venda de algumas marcas do Grupo Impresa à empresa *Trust In News* – como as revistas Visão, Exame e Exame Informática –, que ocorreu em janeiro de 2018, a equipa de quatro profissionais passou, pouco tempo depois, apenas a gerir apenas as páginas do Expresso, da SIC Notícias e da revista Blitz.

Nesta experiência – que durou apenas duas semanas – pude conhecer as estratégias do Grupo Impresa para promover as notícias no seio das redes sociais, as quais mudam de rede para rede. No *Facebook*, as publicações apenas contêm um *link* e uma entrada, e são promovidas em contínuo. No *Twitter*, a periodicidade de promoção é semelhante à do *Facebook*, mas as publicações apenas contêm os títulos e os *links* das notícias. No *Instagram* são publicadas fotografias que marcam a atualidade com frequência e todos os dias é construído um *InstaStories*, que reúne fotografias e os principais tópicos que explicam os temas mais importantes do dia. E no *LinkedIn* são publicados apenas três ou quatro notícias por dia sobre economia e política.

Este processo, o de promoção das notícias nas redes, é atualmente suportado por uma plataforma *online*, o *Swonkie*, que veio facilitar a gestão das redes sociais do grupo. Antes, os profissionais tinham de copiar os *links* e as entradas das notícias do *site* para as redes e tinham também o desafio constante de as manter atualizadas sob curtos períodos de tempo. A página do Expresso deveria ser atualizada, em média, de cinco em cinco minutos no *Facebook*, por exemplo. Com a integração do *Swonkie* esta realidade mudou. Nesta plataforma digital vão surgindo as últimas notícias publicadas nas páginas *web* dos respetivos *media*, as quais vão sendo convertidas automaticamente nos diferentes formatos de publicação para as redes sociais. Isto é, das notícias originalmente publicadas no *site*, o *Swonkie* seleciona, de forma autónoma, as informações necessárias para criar publicações, como, por exemplo, os *links* das mesmas e a entradas, no caso de ser no *Facebook*. E ainda tem a funcionalidade de automatizar o processo de promoção de notícias nas redes e de delinear um intervalo de tempo entre as publicações. Mas mesmo com o trabalho facilitado, tínhamos de continuar atentos para evitar erros e publicações repetidas.

Outra tarefa que executei diariamente passou por selecionar cinco ou seis notícias que deveriam ser agendadas no *Facebook* para serem depois publicadas, entre a meia noite e as 8 horas da manhã – o período de tempo em que não há profissionais na redação. E para fazê-lo analisávamos as notícias mais lidas no *Google Analytics*, bem como as que mais reações provocavam no *Facebook*, em forma de gostos, comentários e partilhas. Esta estratégia também foi a utilizada na hora de escolher notícias para partilhar noutras páginas – por exemplo, um conteúdo da SIC Notícias na página do Expresso. O objetivo era claro: selecionar as notícias que mais atraem o público, para aumentar, ainda mais, o número de visualizações.

Para conhecer os resultados do trabalho desenvolvido ao longo do dia, a equipa recorre também ao *Google Analytics*, o qual regista em tempo real o número de leitores que acedem aos *websites*, nomeadamente do Expresso. E regista também as vias pelas quais o público o acede. O desafio para a equipa era o de manter o número de leitores que visitam o *website* do Expresso via *Facebook* sempre superior ao número de leitores que o consultam via orgânica, isto é, por uma pesquisa no motor de busca da *Google*, por exemplo.

Esta experiência, em particular, permitiu-me conhecer de perto um importante veículo de comunicação do Expresso, as redes sociais. Esta é uma das fortes apostas do Expresso no

universo do *marketing* digital, a qual visa, sobretudo, promover os conteúdos que mais interesse despertam aos leitores. Nesta curta passagem ficou claro para mim o poder das decisões dos diretores do jornal sobre as estratégias de promoção das notícias: as que republicadas são as que mais reações e curiosidade geram, são as que, do ponto de vista financeiro, podem aumentar a rentabilidade da publicidade no seio digital – que, note-se, é tanto maior quanto mais visualizações os conteúdos obtiverem. Embora reconhecendo que a essência e os objetivos deste campo são distintos dos do jornalismo, as estratégias adotadas não têm, na minha perspetiva, de seguir necessariamente este rumo capitalista. A estratégia poderia passar por republicar e destacar as notícias que marcaram o dia, segundo uma perspetiva jornalística, por exemplo.

### ***Rubrica 2:59: O mercado do café português***

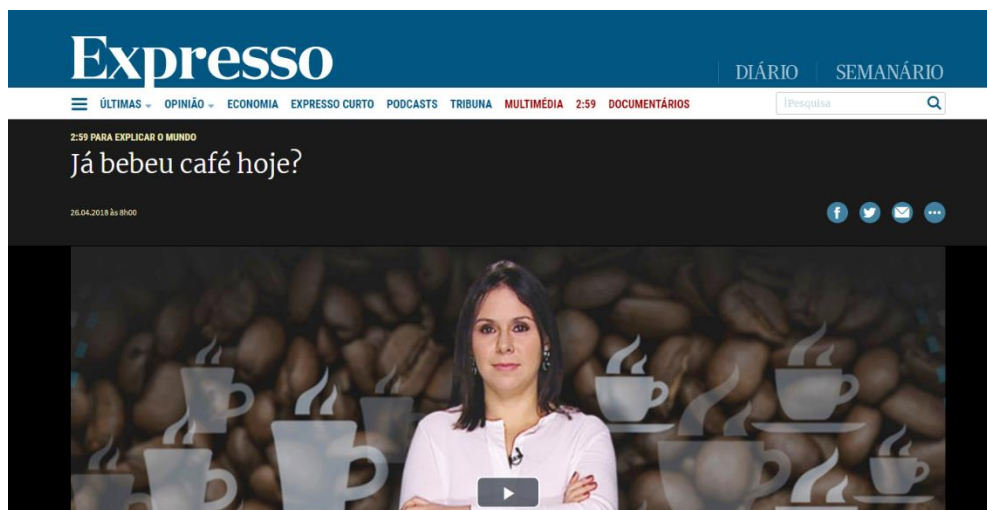
Integrar a rubrica 2:59 foi um dos últimos projetos que abracei – nomeadamente em fevereiro de 2018 – e que marcou fortemente a minha passagem pela vertente digital do Expresso. Para que tal fosse possível, primeiro selecionei um tema de interesse público, passível de ser investigado: o mercado do café português. Apresentei-o à Joana Beleza, coordenadora da rubrica, que logo o aprovou.

Para recolher bases de dados passíveis de serem analisadas, consultei as páginas oficiais das entidades – nacionais e internacionais – que disponibilizam periodicamente informações sobre o mercado do café. Sobre o consumo mundial de café na Organização Internacional do Café (OIC) e sobre o consumo *per capita* na Europa no relatório mais recente da *European Coffee Federation*. Tendo em foco o mercado português, recolhi dados no *Eurostat* sobre os principais países importadores de grãos verdes de café, bem como os países para os quais Portugal exporta grãos de café torrados – os quais foram exportados numa folha de cálculo *Excel* com formato.xlsx. E para compreender as tendências de consumo de café em Portugal solicitei diretamente à *Marktest* os dados mais recentes sobre o assunto.

Perante tantos dados provenientes de fontes distintas, deparei-me com um primeiro desafio: compilar e cruzar dados com diferentes unidades de medida – como milhares de sacas de café com 60 quilogramas (OIC) com milhares de quilogramas de café (*Eurostat*). E depois segui para a fase de análise de dados que recaiu, sobretudo, para um campo comparativo. Tendo por base a análise comparativa, foi possível, por exemplo, perceber que a população europeia é a que mais consome café no mundo; que a Finlândia é o país

do espaço europeu que registou maior consumo de café *per capita* em 2016; que o Vietnã e o Brasil são os países que mais café verde exportam para Portugal; e que Espanha, Grécia e França são os principais destinos do café torrado em território português. Para melhor compreender a evolução do consumo de café mundial e dos consumos entre os continentes, calculei as variações dos respetivos consumos entre 2016 e 2017, e converti-as em frequências relativas. Com este cálculo foi possível perceber, por exemplo, quais foram os continentes que registaram o maior aumento no consumo de café entre esses dois anos. Neste caso, os resultados apontam para um crescimento maior na Ásia e na Oceânia (2,5%) entre 2016 e 2017. E calculando a frequência relativa do consumo europeu face ao mundial, foi possível perceber que os europeus são responsáveis por um terço do consumo global. Para compreender a evolução do consumo de café consoante o local – em casa ou na rua –, também calculei as respetivas variações entre 2013 e 2016. A partir destes resultados foi possível verificar que os portugueses tendem a consumir mais café em casa e menos na rua. Com intuito de encontrar explicações para algumas tendências verificadas – como, por exemplo, o elevado consumo de café na Finlândia – contactei a secretária geral da Associação Industrial e Comercial do Café Cláudia Pimentel – via telefone e via *email*. As suas declarações foram fundamentais para compreender o significado destes e outros dados.

O fio condutor da história que delineei partiu de uma perspetiva geral, do consumo mundial, que logo se afunilou no consumo europeu e posteriormente no português. Escrever o guião deste 2:49 foi, sem dúvida, a maior dificuldade que senti, sobretudo, pelas mudanças existentes no discurso. No registo audiovisual – em que é gravado a rubrica –, o discurso deve ser o mais coloquial possível, as frases simples e curtas para promover uma leitura fácil e fluída. Ler o guião em voz alta à medida que o ia escrevendo foi a estratégia que encontrei para compreender se o discurso estava ou não a ser fluído. Depois de compilar os dados no ficheiro *Excel* e de redigir o guião, enviei o material à jornalista Joana Beleza para que ela o pudesse rever e validar. Foi nos estúdios da SIC que gravei o 2:59, que prometia explicar o mercado de café português, juntamente com a jornalista Cristina Pombo, que gravou a versão que foi publicada no Expresso (diário e *site*) no dia 26 de abril de 2018 – a interface do *site* está representada na figura 4.5. Com o apoio do realizador André Atayde foi possível, depois de várias tentativas, gravar uma versão minha passível de ser depois editada pelo *designer* gráfico Carlos Paes.



**Figura 4.5** – Interface do *site* do 2:59 para explicar o mercado do café português

**Fonte:** Cópia Integral de Pombo, Sousa, Atayde, Beleza, & Paes (2018)

Esta experiência mostrou-me, em particular, a importância de ser versátil para escrever para qualquer formato, para fazer jornalismo multiplataforma, e evidenciou também a importância de existir convergência entre meios de comunicação no seio de um grupo mediático, pois se não existisse a parceria entre o Expresso e a SIC a rubrica 2:59 não seria, provavelmente, gravada com a mesma qualidade. Esta experiência não só enriqueceu o meu estágio, como também marcou o próprio Expresso. Segundo a coordenadora multimédia Joana Beleza, fui a primeira estagiária a gravar um 2:59, e, portanto, a fazer parte de todo o processo: desde a recolha de dados e redação do guião à gravação e edição do mesmo.

#### **4.2.3 Desafios e Aprendizagens**

Foram várias as vezes em que tomei a iniciativa de apresentar aos editores alguns temas passíveis de serem aprofundados através da exploração dos dados mais recentes disponibilizados pelas entidades oficiais. E também foram várias as vezes em que fui desafiada pelos editores a investigar dados sobre os temas quentes na agenda mediática.

Nem todas as propostas foram avante. Nem todos os temas apresentavam uma novidade. Para distinguir o que interessa do que não interessa, o que são informações novas e de interesse público e o que não são, o acompanhamento diário da jornalista Raquel Albuquerque, bem como dos editores Germano Oliveira e José Cardoso foi fundamental. Estes profissionais ensinaram-me a refletir sobre os temas e a colocar-me no lugar de quem lê. E esta aprendizagem, que despertou em mim a presença de sentido crítico – referenciada por Kayse (2012) como um dos aspetos que os jornalistas de dados devem



estimular – é apenas uma das que guardo comigo. Outra, a capacidade de investigar persistentemente, levou-me mais longe. Foi a perseverança que me levou a encontrar novas perspetivas nos dados, tendências ainda não evidenciadas pela imprensa sobre temas de interesse público.

A firmeza na busca constante por informações novas traduziu-se em várias horas, dias e até semanas de trabalho. Foram várias histórias baseadas em dados – como é o caso das mulheres autarcas – que desenvolvi sob longos períodos de tempo. E são várias as razões que o justificam. Por vezes, encontrar bases de dados que possuam exatamente a informação que procuramos nem sempre é possível. A história que desenvolvi sobre a idade mínima para o consentimento sexual é um exemplo. A escala do trabalho foi reduzida à realidade dos países do espaço europeu, porque, após várias horas de pesquisa, a única base de dados que encontrei que reunia dados sobre todos os países do mundo – a ideia inicial do trabalho, proposta pelo editor Germano Oliveira – não apresentava traços de credibilidade. Este caso, em concreto, mostrou-me a importância de refletir sobre os dados que recolhemos e, portanto, de verificar a sua credibilidade.

Outro aspeto que explica o imenso tempo despendido na concretização de trabalhos deste tipo tem que ver com o processo de gestão de dados. Não encontrando mecanismos automáticos para limpar os dados, este processo pode ser demorado. E, por vezes, encontrar a melhor perspetiva para cruzar e analisar os dados nem sempre é fácil. Também é necessário ter uma atenção redobrada para interpretá-los, para traduzir os dados em ideias claras, percetíveis e, sobretudo, corretas. Além do referido, conseguir entrevistar especialistas rapidamente nem sempre é possível. Por vezes, demoram semanas a responder a uma entrevista por escrito enviada por *email*. Um outro aspeto que considero importante referir está relacionado com a gestão de tempo e de tarefas, um desafio encarado diariamente pelos jornalistas. Por vezes, tinha simplesmente de largar estes trabalhos para me dedicar a outros, temas quentes que, de repente, surgiam na agenda mediática. Em maior parte dos casos desenvolvia as histórias sob pressões de tempo, para que pudessem ser publicadas às seis da tarde nos temas principais do Expresso Diário. Todas estas dificuldades, que surgem nas diferentes fases do processo de jornalismo de dados, são os principais motivos que explicam as várias semanas e até meses de dedicação estes trabalhos baseados em dados.

#### **4.2.4 Principais tópicos de investigação**

##### ***As principais fontes de informação acedidas***

No decorrer do estágio, fui procurando na *web* bases de dados com interesse passíveis de serem investigadas, sobretudo, nos *websites* oficiais das organizações não - governamentais internacionais e nacionais (como as Nações Unidas e o PorData), das organizações governamentais internacionais e nacionais (como o *Eurostat* e o Instituto Nacional de Estatística) e nos meios de comunicação social, sobretudo internacionais, como o *The Guardian* e o *ProPublica*. Estas foram as principais fontes de informação a que recorri para obter bases de dados. As páginas *web* das associações e o arquivo digital do Expresso, gerido pela Gesco, também foram fontes a que acedi, mas com uma frequência menor.

A título de exemplo, considero importante referir que para desenvolver os temas sobre a mutilação genital feminina, o casamento infantil e a idade mínima para consentir um ato sexual, descarreguei bases de dados sobre os referidos temas do *website* das Nações Unidas. Para investigar o risco de pobreza no espaço europeu, uma matéria que me despertou especial interesse em aprofundar, recorri às bases de dados mais recentes do *Eurostat*. E a história que desenvolvi sobre os tiroteios em massa nos Estados Unidos da América foi assente na base de dados compilada e disponibilizada *online* pela revista *Mother Jones*.

##### ***Bases de dados internas do Expresso***

O Expresso possui um sistema de arquivo interno, gerido pela Gesco, também inserido no Grupo Impresa. Este arquivo é acedido pelos profissionais da redação através do sistema denominado *Elvis*, onde é possível encontrar milhões de documentos e fotografias, que todos os dias são gerados pelas marcas do grupo. E também é possível consultar a bases de dados, que são criadas pelos jornalistas e infografistas não de forma sistemática, mas esporadicamente. Estes profissionais fazem-no quando, por exemplo, sentem a necessidade de organizar um conjunto de dados e informações contidas em relatórios técnicos com centenas de páginas. Não o fazem sistematicamente, porque, segundo o que apurei no decorrer do estágio, não existem recursos humanos somente dedicados à prática de jornalismo de dados e, portanto, somente focados em pesquisar, recolher e tratar bases de dados.

As informações, bases de dados e fotografias criadas pelas marcas do Grupo Impresa podem ser fornecidas pela Gesco a outras organizações mediante um pagamento acordado entre ambas as partes. Esta viabilização dos conteúdos é outra forma de o grupo gerar receitas.

### ***Os principais sistemas de informação utilizados***

São vários os sistemas de informação utilizados no seio da redação do Expresso. Para comunicar internamente, por exemplo, os profissionais usam com frequência o sistema VOIP (*Voice Over Internet Protocol*), que é essencialmente utilizado para enviar mensagens rápidas e para telefonar para os colegas na redação. O processo de jornalismo de dados, em particular, é suportado pelos vários sistemas de informação disponíveis na redação. A tabela 4.2 que se segue procura, então, indicar quais são os principais sistemas utilizados em cada fase do processo de jornalismo de dados, fazendo corresponder os principais utilizadores, bem como as suas funções.

**Tabela 4.2** – Sistemas de informação utilizados no processo de jornalismo de dados

**Fonte:** Elaboração própria

<b>Fase do processo de jornalismo de dados</b>	<b>Sistema de informação</b>	<b>Função</b>	<b>Utilizadores</b>
Investigação de temas de interesse público	<i>Google Analytics</i>	Reúne dados sobre os comportamentos dos leitores: número de visualizações, tempo de leitura, entre outras.	Editores, jornalistas, gestores das redes sociais
Recolha de dados e informações	<i>Elvis - Gesco</i>	Sistema de arquivo do Grupo Impresa, no qual estão armazenados todos as notícias redigidas, todas as bases de dados criadas, todas as fotografias captadas, áudios, etc.	Editores, jornalistas, infografistas, fotógrafos
Tratamento, análise e interpretação de dados	<i>Microsoft Office Excel</i>	É o principal sistema utilizado para trabalhar os dados, proceder a cálculos, organizar os dados em tabelas e visualizá-los em gráficos	Editores, jornalistas, infografistas
	<i>SPSS-Statistics</i>	Sistema de tratamento e cálculo estatístico	Jornalistas, infografistas
Redação da notícia	<i>Microsoft Office Word</i>	Sistema utilizado para compilar, organizar informação e redigir notícias	Editores, jornalistas, infografistas

<b>Fase do processo de jornalismo de dados</b>	<b>Sistema de informação</b>	<b>Função</b>	<b>Utilizadores</b>
Construção de Visualizações	<i>Tableau Software</i>	<i>Software de business intelligence</i> , que permite analisar dados de forma rápida e construir visualizações com os mesmos, sobretudo interativas	Infografistas
Publicação da notícia	<i>Content Station</i>	Sistema utilizado para compilar vários elementos das notícias (texto, gráficos, imagens, etc.) e para os publicar no Expresso Diário. Este é também o sistema utilizado para introduzir os conteúdos do semanário e para geri-los	Editores, jornalistas, infografistas, fotógrafos
	<i>Backoffice</i>	Sistema utilizado para compilar vários elementos das notícias e para os publicar no <i>website</i> do Expresso	Editores, jornalistas, infografistas, fotógrafos

### ***Otimização de processos produtivos***

Perante a instabilidade do noticiário do Expresso, a redação adota estratégias de distribuição de conteúdos pelas diferentes plataformas, as quais lhe permitem tornar o processo de gestão do noticiário flexível. Se uma notícia redigida para o semanário, por exemplo, for “queimada” por outros sites – uma expressão que retrata a perda da exclusividade do conteúdo –, o Expresso publica-a, de imediato, no *site*.

Este processo, o de transplantar os conteúdos para diferentes plataformas, é necessário, mas não está facilitado, requer tempo, e há pelo menos um jornalista encarregue desta tarefa. Dado que as plataformas são geridas em sistemas diferentes, sem conexão entre si, as notícias têm de ser literalmente copiadas de um sistema para o outro. Do *Content Station* para o *Backoffice* se um tema passar do papel para o *site*, no primeiro exemplo. E isso significa que todo o conteúdo (o título, a entrada, o corpo da notícia, as fotografias, gráficos, assinatura, entre outros elementos possíveis) tem de ser literalmente copiado de um sistema para o outro.

Este processo consome tempo que poderia ser poupado se, por exemplo, for instalado um sistema integrado de gestão na redação do Expresso. Este sistema – que possui uma única

base de dados e que procura integrar a gestão das diferentes plataformas do Expresso – faria toda a diferença na redação, pois nele seria possível inserir todo o conteúdo de uma notícia uma única vez e indicar qual a plataforma de publicação do jornal: *site*, Expresso Diário ou semanário. Posto isto, considero que uma aposta num sistema integrado de gestão no jornal Expresso poderia otimizar o processo de publicação de notícias nas suas diferentes plataformas.

#### **4.2.5 Conclusão do Estágio**

Mesmo no meu último dia na redação – dia 18 de março de 2018 – surpreendi e deixei-me surpreender. Para adoçar a despedida, levei um cestinho de amêndoas que coloquei junto da máquina de café. E, com o mesmo intuito, a jornalista Manuela Goucha Soares levou um bolo da Padaria Portuguesa. Fiz questão de partilhar o momento com todos os editores, jornalistas, fotojornalistas e infografistas que me acompanharam durante seis meses e que me ajudaram a crescer enquanto futura profissional. Este lanche, de certa forma, assemelhou-se aos pequenos banquetes que muitas vezes surgiam na redação: quer para festejar um aniversário, quer para desejar boa sorte a alguém que está de partida.

Cada uma das experiências que referi moldou a forma como olho para os acontecimentos e como faço jornalismo e, em especial, jornalismo de dados. Cada uma contribuiu, de forma particular, para o meu desenvolvimento enquanto futura jornalista. Nesta aventura, que foi o meu estágio no Expresso Digital, aprendi o valor da versatilidade, da iniciativa e, principalmente, do trabalho em equipa. Sem o apoio de todos os profissionais que mencionei ao longo deste diário, esta experiência não teria sido tão rica em aprendizagens.

## **5 Investir e inovar em jornalismo de dados: o caso do jornal Expresso**

O jornalismo de dados tem sido de tal forma adotado no seio dos meios de comunicação pelo mundo fora que já é considerado uma nova tendência jornalística internacional (Coddington, 2015). E não esqueçamos que a sua adoção gera mudanças não só nas redações e nas rotinas dos jornalistas que a ele se dedicam, como também nos leitores. Algumas dessas mudanças já foram referenciadas no primeiro capítulo, que aborda o estado da arte do jornalismo de dados. Outras ainda estão por descobrir. Aquelas que apenas se podem conhecer quando no meio entramos, quando de perto observamos a integração do jornalismo de dados numa redação e quando interrogamos os envolvidos.

Face a esta necessidade de conhecer mais sobre este fenómeno – a integração do jornalismo de dados nas redações e o seu impacto nos leitores – estudar profundamente um caso, um meio de comunicação que abraça o jornalismo de dados, é o caminho. E, por esse motivo, o caso do jornal Expresso é o eleito. Esta estratégia de pesquisa, de acordo com Prodanov & Freitas (2013), enquadra-se na definida como estudo de caso, pois propõe conhecer uma unidade, um fenómeno específico, partindo de um caso passível de o ilustrar. Para estudar como os jornalistas do Expresso integram o jornalismo de dados nas suas rotinas e para conhecer a perceção dos seus leitores, há que aplicar diferentes técnicas de recolha de dados direccionadas a estes universos de pesquisa.

Uma dessas técnicas diz respeito à observação participante, a qual permitiu não só ter uma perceção da integração desta prática na redação do Expresso, como também vivenciá-la. Esta experiência, por si só, já foi uma aposta ganha no que diz respeito ao estudo das rotinas produtivas dos jornalistas, pois permitiu conhecer de perto todas as fases, todos os entraves que podem surgir, todas as nuances. Mas para conhecer como os jornalistas do Expresso aplicam no seu dia a dia os processos inerentes ao jornalismo de dados, questionar cada um deles sobre o assunto afigura-se na estratégia mais eficaz. E esta também foi a técnica de recolha de informação utilizada para conhecer de perto a perceção dos leitores – e potenciais leitores do jornal – quanto às histórias baseadas em dados.

## **5.1 Como é que os jornalistas do Expresso abraçam o jornalismo de dados?**

### **5.1.1 Recolha, seleção e métodos de análise**

Para conhecer de que forma os jornalistas do Expresso aplicam o processo de jornalismo de dados no seu dia a dia, a investigadora construiu um questionário – utilizando para tal a aplicação “Formulários do Google” – que foi disponibilizado entre 26 de fevereiro e 18 de março do presente ano. Dos 56 jornalistas do Expresso, que receberam o questionário no *email*, apenas 24 responderam, sendo, por isso, a taxa efetiva de resposta de 42,9%.

Dividido em três secções distintas, o questionário procurou recolher os dados sociodemográficos dos jornalistas, bem como os que permitem definir um perfil profissional genérico. E procurou também compreender, na terceira secção, vários aspetos relacionados com a integração do processo de jornalismo de dados nas rotinas produtivas dos jornalistas. O questionário foi elaborado tendo por base os conhecimentos adquiridos no decorrer do estágio e foi desenhado de forma a obter um padrão de respostas, pelo que a grande maioria das questões são fechadas. Numa ou noutra questão foi adicionada uma resposta alternativa – nas quais a investigadora considerou que poderia haver mais opções a serem mencionadas pelos jornalistas.

Em termos gerais, a análise que se segue procura confrontar duas perspetivas sobre o jornalismo de dados: a visão dos jornalistas mais experientes, que iniciaram a sua carreira no século XX, com a dos jovens jornalistas, que deram os primeiros passos na profissão no presente milénio. A viragem do milénio foi o ponto temporal delineado para separar estas duas gerações de jornalistas por uma razão: as fortes mudanças ocorridas. Esta passagem trouxe consigo uma avalanche de evoluções tecnológicas que despoletaram o desenvolvimento do jornalismo digital e, em especial, do jornalismo de dados. Os jornalistas que iniciaram a carreira antes do ano 2000, que marca esta passagem, vivenciaram todas as mudanças desencadeadas, beneficiando de um clima de expansão (Rebelo, 2014). Conheceram a redação do Expresso antes da vertente digital ganhar força nas redações, antes de existirem bases de dados e *software* acessíveis a todos, antes do jornalismo de dados se tornar numa tendência jornalística internacional. Conheceram, possivelmente, outras práticas de trabalhar com dados, como é o caso do jornalismo de precisão de Meyer. Já o outro grupo de jornalistas do Expresso, que dá início à sua atividade profissional depois de 2000, conhece uma realidade jornalística diferente, marcada por uma crise económica e financeira que não só diminuiu postos de trabalho

como também deteriorou as condições laborais dos jornalistas (Rebelo, 2014). Quando estes jornalistas iniciaram a sua atividade profissional no Expresso, a vertente digital já fazia parte da redação e os dados e o *software* já eram, de certo modo, acessíveis ao público. Estes jornalistas conheceram, portanto, um contexto profissional diferente dos que iniciaram a carreira ainda no século XX. E tendo em conta estas diferenças, torna-se relevante confrontar as suas perspetivas para perceber se estas duas gerações distintas de jornalistas do Expresso encaram o jornalismo de dados de modo semelhante ou diferente. Um dos pontos de análise visa descobrir, por exemplo, se há ou não diferenças nas dificuldades sentidas pelos jornalistas quando executam tarefas que se inserem no processo. Se a frequência de produção de histórias baseadas em dados é ou não a mesma. Se utilizam os mesmos sistemas de informação para tratar os dados e se recorrem às mesmas fontes de informação. Se encaram o investimento nesta forma de fazer jornalismo do mesmo modo. Posto isto, cruzar as respostas dos jornalistas com largos anos de carreira com as dos menos experientes poderá ser um caminho a seguir para conhecer como duas gerações distintas vivem e encaram o universo do jornalismo de dados.

Para realizar a referida análise, os dados provenientes das respostas dos jornalistas foram tratados e manipulados no sistema *Microsoft Office Excel*. As técnicas utilizadas passam, sobretudo, pela contabilização do número de repostas (frequência absoluta) e pelo cálculo da frequência relativa de cada variável. A interpretação terá por base a comparação dos resultados entre as duas gerações definidas. E os gráficos e tabelas que os representam foram também desenhados no referido sistema.

Assim, a partir da análise dos resultados obtidos será possível compreender se diferentes gerações de jornalistas do Expresso encaram e integram o jornalismo de dados nas suas rotinas de modo similar ou de modo diferente.



### **5.1.2 Caracterização sociodemográfica e profissional das amostras**

Numa análise inicial às respostas dos 24 jornalistas do semanário Expresso, foi possível verificar que exatamente metade dos jornalistas inquiridos começou a sua carreira antes da viragem do milénio. E a outra metade iniciou-a na primeira ou na segunda décadas do século XXI. Para conhecer um pouco mais das características destas duas amostras definidas pela investigadora, há que examinar os dados que permitem traçar um perfil sociodemográfico e profissional genérico destes dois grupos de jornalistas do Expresso.

Os jornalistas que possuem mais anos de experiência – o primeiro grupo definido – iniciaram as suas carreiras em três décadas distintas. Exatamente metade iniciou nos anos 90 e um terço começou a carreira de jornalista nos anos 80. Apenas dois jornalistas (o equivalente a 16,7% da amostra) deram os primeiros passos nesta profissão nos anos 70 do século passado. Os dados mostram também que grande fatia da experiência destes jornalistas foi adquirida no jornal Expresso, já que a maioria (66,7%) diz exercer funções no semanário há mais de 20 anos. Exatamente um quarto dos jornalistas trabalha nesta redação há mais de dez anos e menos de 20 e apenas 8,3% entre cinco e dez anos. Estes jornalistas são sobretudo mulheres (75%) e todos possuem idades entre os 36 e os 65 anos. A grande maioria é licenciado (75%), cerca de 8,3% desta geração possui o grau de mestre e 16,7% concluiu o ensino secundário.

No segundo grupo, cerca de 41,7% dos jornalistas iniciaram a carreira na primeira década deste milénio e a maioria (58,3%) iniciou-a na segunda. O jornal Expresso também tem sido uma verdadeira escola para estes jornalistas, já que grande parte diz trabalhar no Expresso há menos de 5 anos (66,7%) e 8,3% entre 5 e 10 anos. Neste grupo, há também alguns jornalistas que trabalham no Expresso entre 11 e 20 anos (16,7%) e apenas um que diz trabalhar há mais de 20 – que talvez tenha exercido outras funções no jornal. Também três em cada quatro jornalistas são mulheres. E no que diz respeito à idade dos jornalistas, o panorama – como seria de esperar – rejuvenesce: um em cada quatro possui menos de 25 anos e metade dos jornalistas deste grupo possuem entre 25 e 35 anos. Ainda assim, um em cada quatro destes jornalistas possui mais de 35 anos e menos de 65. Ao nível de habilitações literárias também há diferenças: exatamente metade desta geração possui mestrado e 41,7% concluiu licenciatura. Apenas um dos jornalistas admite ainda frequentar a faculdade (o equivalente a 8,3%) e nenhum diz ter concluído apenas o ensino secundário.

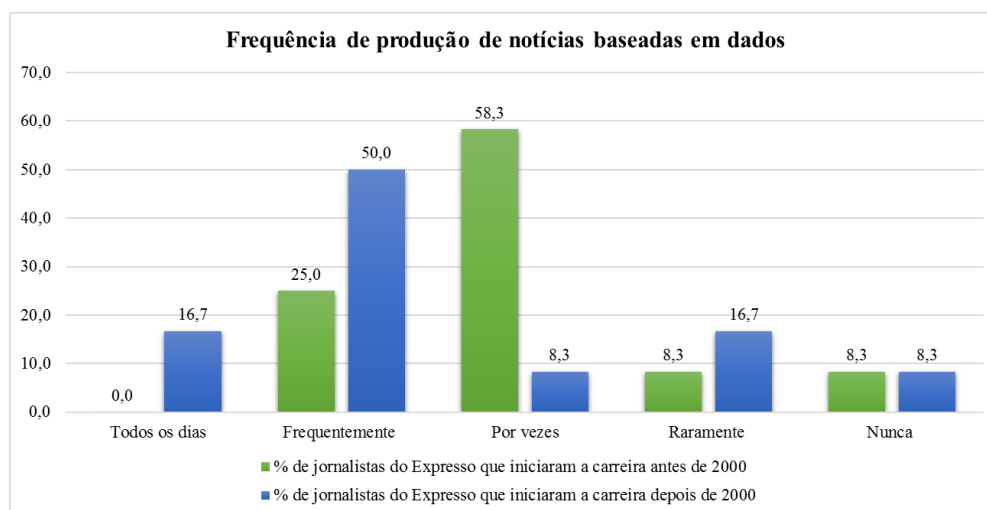
### 5.1.3 O jornalismo de dados aos olhos de duas gerações distintas

Caminhando sob vários aspetos inerentes à prática do jornalismo de dados – tendo por base a literatura – a presente análise irá reter o que dizem estes dois olhares: a perspetiva de quem já tem pelo menos dezoito anos de experiência como jornalista, e a visão de quem iniciou a carreira no presente século. O objetivo é claro: perceber se há diferenças e semelhanças entre os vários aspetos em análise. Em suma, comparar os olhares destas duas gerações tendo como pano de fundo o jornalismo de dados é o foco da presente análise.

Será que os jornalistas com largos anos de carreira escrevem mais ou menos histórias baseadas em dados do que os jornalistas menos experientes? Esta é uma questão de partida que logo elucida, que logo revela se este fluxo de trabalho é ou não inserido nas rotinas produtivas destes dois grupos de jornalistas.

**Gráfico 5.1** – Frequência de produção de notícias baseadas em dados pelos jornalistas que iniciaram a carreira antes e depois do ano 2000

Fonte: Elaboração própria



Os dados, apresentados no gráfico 5.1, mostram que apenas um em cada quatro jornalistas que compõem o primeiro grupo, o dos mais experientes, produzem notícias baseadas em dados com frequência ou todos os dias. Mais de metade diz fazê-lo apenas por vezes (58,3%) e cerca de 16,7% raramente ou nunca o faz. Já a maioria dos jornalistas menos experientes (66,7%) diz construir notícias baseadas em dados com frequência ou todos os dias. E ainda há os que – neste grupo – as produzem algumas vezes (8,3%) e os que raramente ou nunca o fazem (25%). Estes resultados indicam – ou podem indicar – que os jornalistas com menos anos de carreira abraçam com mais frequência o jornalismo de dados do que os jornalistas mais experientes.

No que diz respeito às várias fases inerentes ao processo do jornalismo de dados – referidas na literatura –, os jornalistas foram levados a refletir sobre as tarefas que executam com mais e menos dificuldade. As perceções dos jornalistas sobre estes aspetos que iniciaram a carreira no século passado e no presente milénio encontram-se reunidas na tabela 5.1, que se segue.

**Tabela 5.1** – Perceção do grau de dificuldade dos jornalistas mais e menos experientes do jornal Expresso relativamente às tarefas incluídas no processo de jornalismo de dados (em % face ao total de respostas de cada grupo)

**Fonte:** Elaboração própria

Grau de dificuldade	Difícil		Dificuldade Intermédia		Fácil	
	Antes 2000 (%)	Depois 2000 (%)	Antes 2000 (%)	Depois 2000 (%)	Antes 2000 (%)	Depois 2000 (%)
Pesquisar informação	0	0	33,3	25,0	66,7	75,0
Recolha de bases de dados / relatórios técnicos	25,0	8,3	41,7	41,7	33,3	50,0
Análise de dados (limpeza, cruzamento, cálculo)	33,3	25,0	58,3	41,7	8,3	33,3
Interpretação de dados	0	8,3	58,3	33,3	41,7	58,3
Extrair informação dos dados	16,7	0	33,3	58,3	50,0	41,7
Selecionar informação dos dados	8,3	0	41,7	50,0	50,0	50,0
Recolher declarações de especialistas	0	0	50,0	41,7	50,0	58,3
Redação de histórias baseadas em dados	0	0	33,3	33,3	66,7	66,7
Solicitação e seleção de conteúdos multimédia	25	16,7	50,0	33,3	25,0	50,0

Observando os pontos em comum dos resultados obtidos, verifica-se que há apenas duas tarefas consideradas fáceis pela maioria dos jornalistas de ambos os grupos: uma que marca o início do processo, a pesquisa de informação (considerada fácil por cerca de 66,7% dos jornalistas experientes e por 75% dos jornalistas menos experientes) e a outra, que faz parte do final do mesmo, a redação de histórias baseadas em dados. Esta última é acessível para cerca de 66,7% dos jornalistas de cada grupo e revela alguns obstáculos para um terço destes profissionais. E há ainda uma tarefa que é executada com facilidade por exatamente metade dos profissionais de ambas as gerações e que, ao mesmo tempo, revela pelo menos algumas dificuldades para a outra metade: selecionar informação dos dados. Outro ponto em comum diz respeito à tarefa que revela mais dificuldades para estas duas gerações de jornalistas: a análise de dados, uma fase específica do processo de jornalismo de dados. A maioria dos jornalistas mais experientes (58,3%) admite sentir algumas dificuldades em analisar dados e um terço considera-a mesmo uma fase difícil.

Esta tarefa é executada com facilidade por apenas um jornalista com mais de 18 anos de carreira (8,3%). No grupo dos jornalistas menos experientes, um em cada quatro diz que analisar dados é árduo e cerca de 41,7% deste grupo admite que existem alguns obstáculos. Por outro lado, um terço destes jovens jornalistas analisam dados – isto é, limpam, cruzam e realizam cálculos – com facilidade. Os resultados evidenciam que a maioria dos jornalistas de ambas das gerações sentem pelo menos algumas dificuldades quando analisam dados.

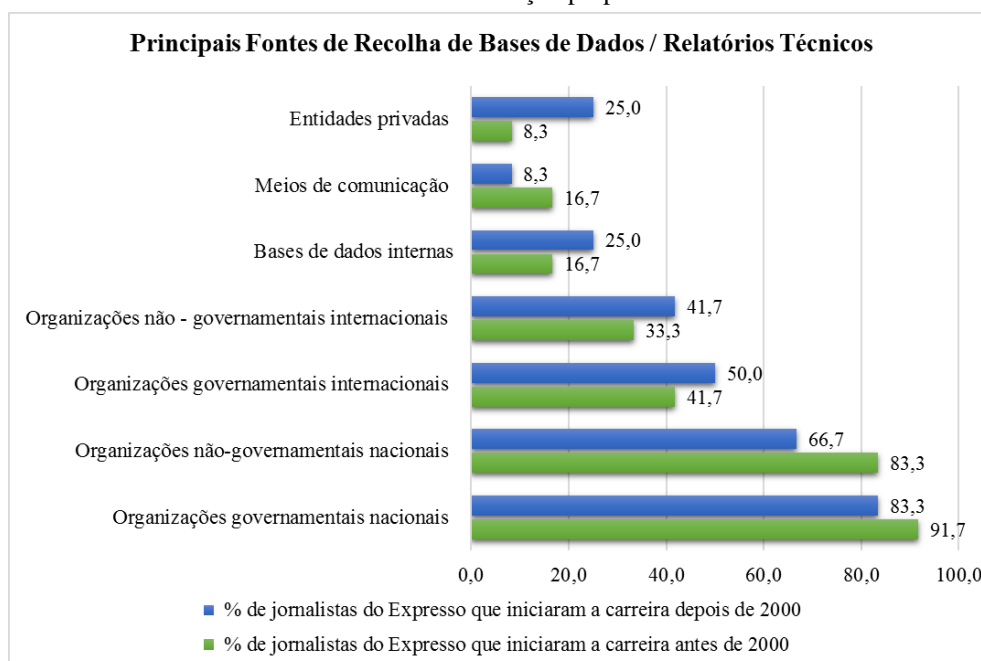
As restantes cinco tarefas apresentam diferenças no grau de dificuldade entre gerações. A recolha de dados é uma delas e divide opiniões no grupo dos jornalistas menos experientes: metade considera fácil e a outra metade diz sentir pelo menos algumas dificuldades. Já no outro grupo, dos jornalistas com mais anos de carreira, há uma tendência clara: a maioria dos jornalistas experientes (cerca de 66,7%) diz sentir pelo menos alguns entraves na sua concretização e apenas um terço considera que recolher bases de dados é acessível. A interpretação de dados também apresenta diferenças entre grupos. Esta é uma tarefa fácil para a maioria dos jovens jornalistas (58,3%) e difícil para a maioria dos experientes (58,3%). E, inversamente, extrair informação dos dados apresenta pelo menos alguns obstáculos a maioria dos jovens jornalistas (58,3%) e divide opiniões no grupo dos mais experientes: metade considera fácil e a outra metade diz sentir alguns entraves na sua execução. Esta geração de jornalistas divide-se do mesmo modo, quando o assunto é recolher declarações de especialistas. Já para a maioria dos jornalistas menos experientes (58,3%) esta é uma fase acessível. Quando levados a refletir sobre o grau de dificuldade de solicitar e selecionar conteúdos multimédia, três em quatro jornalistas mais experientes revelam sentir pelo menos alguns obstáculos. E no grupo dos jovens jornalistas as perceções voltam a dividir-se pela metade, sendo que tanto há jornalistas desta geração que consideram esta uma tarefa fácil, como também os que possuem algumas dificuldades em executá-la. Posto isto, verifica-se que há quatro tarefas que possuem um grau de dificuldade semelhante para ambas gerações e há cinco que possuem um grau distinto.

Um dos primeiros passos do processo de jornalismo de dados é a recolha de bases de dados ou de relatórios técnicos, a principal matéria prima utilizada pelos jornalistas para construir histórias que nascem dos dados (Coddington, 2015). Por este motivo, um aspeto abordado no questionário procura compreender, precisamente, de onde vêm os dados utilizados pelos jornalistas do Expresso para construir histórias. Neste sentido, a análise

que se segue procura decifrar quais são, portanto, as principais fontes de informação acedidas pelos dois grupos de jornalistas para recolher dados, as quais estão sintetizadas no gráfico 5.2.

**Gráfico 5.2** – Principais fontes de recolha de bases de dados ou relatórios técnicos utilizadas pelos jornalistas do Expresso que iniciaram a carreira antes e depois do ano 2000 (em % face ao total de respostas de cada grupo)

Fonte: Elaboração própria



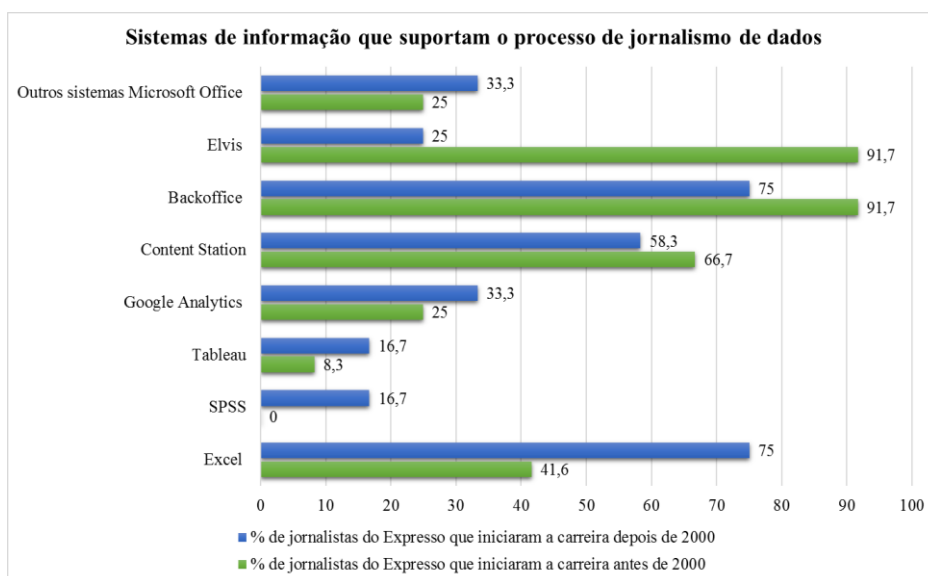
Os resultados do questionário mostram uma tendência evidente: há um consenso entre gerações. As fontes mais acedidas pelos jornalistas que iniciaram a carreira antes do ano 2000 são também as mais acedidas pelos jornalistas mais jovens. A generalidade destes dois grupos – 91,7% dos jornalistas mais experientes e 83,3% dos menos experientes – recolhe dados e relatórios técnicos sobretudo de organizações governamentais nacionais, como é exemplo o Instituto Nacional de Estatística. A maioria dos jornalistas dos dois grupos – 83,3% dos mais velhos e 66,7% dos mais jovens – também utiliza dados de organizações não-governamentais nacionais, como é o caso do PorData. Os dados gerados pelos dois tipos de organizações nacionais são, portanto, os mais recolhidos pelos jornalistas do Expresso, independentemente da geração a que pertencem. Em terceiro e quarto lugar, estão as organizações governamentais e não-governamentais internacionais. Metade da geração de jornalistas mais jovens e cerca de 41,7% dos jornalistas mais experientes dizem recolher dados de organizações governamentais internacionais (como é o caso do Eurostat). As organizações não-governamentais internacionais, como as Nações Unidas são acedidas por 41,7% dos jornalistas menos experientes e por um terço

dos jornalistas com mais anos de carreira. Quanto às fontes menos acedidas, apenas 8,3% dos jovens jornalistas e 16,7% dos mais experientes recorrem às bases de dados disponibilizadas pelos meios de comunicação social. As entidades privadas são o tipo de fonte de informação menos acedido pelos jornalistas mais experientes: apenas 8,3% diz recolher dados a partir destas entidades. Já 25% dos jovens jornalistas recolhem dados das entidades privadas, como é o caso da *Maktest*. As bases de dados disponíveis a nível interno são utilizadas também por um em cada quatro jornalistas mais jovens e por cerca de 16,7% dos mais velhos. Sendo esta uma questão que inclui a possibilidade de acrescentar respostas alternativas, é de salientar que um jornalista experiente adicionou uma outra fonte relevante: as associações setoriais.

Um outro aspeto relevante são os sistemas de informação, que suportam o processo de jornalismo de dados. De acordo com Henninger (2013), são estes sistemas que facilitam a gestão e análise de grandes quantidades de dados. Tendo em conta os principais sistemas utilizados no seio do jornal Expresso – percebidos no decorrer do estágio –, a seguinte análise propõe compreender quais são os mais e menos utilizados pelas duas gerações de jornalistas, elucidando a relevância de cada um nas diferentes fases do processo.

**Gráfico 5.3** – Sistemas de informação que suportam o processo de jornalismo de dados utilizados pelos jornalistas do Expresso que iniciaram a carreira antes e depois de 2000

**Fonte:** Elaboração própria



Um dos primeiros passos do jornalismo de dados é – ou pode ser – conhecer o interesse do público sobre um determinado tema passível de ser investigado. E para tal os jornalistas do Expresso podem recorrer às estatísticas disponibilizados pelo *Google*

*Analytics*, que mostram, por exemplo, as histórias já publicadas sobre esse mesmo tema que mais e menos visualizações conquistaram. Tendo por base os resultados do questionário, reunidos no gráfico 5.3, verifica-se que esta ferramenta não é muito utilizada pelos jornalistas do Expresso. Ainda assim, há ligeiramente mais jovens jornalistas a utilizá-la (33,3%) do que os mais experientes (25%). Outra fase que se insere no início do processo de jornalismo de dados diz respeito à pesquisa e recolha de dados, como já foi anteriormente referido. Os resultados do questionário mostram que a generalidade dos jornalistas mais experientes do Expresso – cerca de 91,7% – para além de consultarem as fontes externas já referidas, também recorrem ao *Elvis*, o sistema interno de arquivo do Grupo Impresa. Já os jornalistas mais jovens não parecem ter o mesmo hábito: apenas um em cada quatro admite consultar este sistema.

Recolhidos os dados, segue-se a fase de limpeza, tratamento e análise dos mesmos, a qual é igualmente suportada por sistemas de informação. Na redação do Expresso, o principal sistema disponibilizado aos jornalistas para o fazer é o *Microsoft Office Excel*, o qual é utilizado por três em cada quatro jovens jornalistas e por menos de metade dos mais experientes (41,6%). Outro sistema disponível que auxilia este processo é o *SPSS – Statistics*, o qual é utilizado apenas por 16,7% dos jornalistas que iniciaram a carreira depois do ano 2000 e por nenhum jornalista com vasta experiência profissional. Os dados indicam, assim, que possivelmente há mais jovens jornalistas a utilizar sistemas de análise de dados do que jornalistas mais experientes.

Na fase de construção das histórias baseadas em dados, que inclui a compilação de diferentes conteúdos multimédia, há especialmente dois sistemas de informação que são utilizados pela generalidade dos jornalistas do Expresso, de ambos os grupos: o *Content Station* e o *Backoffice*. Nesta fase também se podem incluir os outros sistemas da *Microsoft Office*, onde se insere o *Word*, pois há jornalistas que têm preferência em redigir histórias nestes sistemas, como é o caso de um terço dos jornalistas mais experientes e 25% dos mais jovens.

No que diz respeito à construção de visualizações, uma das últimas fases do processo de jornalismo de dados, o *Tableau* foi o sistema mencionado no questionário – o qual é também referenciado por Henninger (2013). Ainda assim, este é apenas utilizado por cerca de 16,7% dos jornalistas menos experientes e por cerca de 8,3% dos que somam mais anos de carreira. Estes valores baixos podem estar relacionados com o facto de no Expresso existir uma equipa destinada à produção de visualizações estáticas e interativas,

o que significa que há poucos profissionais a utilizar este tipo de ferramentas. Sendo também esta uma questão que permitia incluir outros sistemas não mencionados, há que referir que um dos jornalistas que iniciou a carreira depois do ano 2000 utiliza o *Indesign* e outro, do mesmo grupo, acrescentou mais dois: o *Carto DB* e o *Adobe Illustrator*.

Um ponto relevante associado ao jornalismo de dados diz respeito às várias capacidades que um jornalista deve desenvolver para o conseguir introduzir nas suas rotinas produtivas. Será que os jornalistas do Expresso consideram que um jornalista de dados deve desenvolver as principais habilidades mencionadas na literatura? Esta é a questão de partida, que visa conhecer as perspetivas dos jornalistas do Expresso – dos mais jovens e dos mais experientes – sobre este tema, as quais estão reunidas na tabela 5.2.

**Tabela 5.2** – As capacidades que os jornalistas de dados devem desenvolver na perspetiva dos jornalistas do Expresso que iniciaram a carreira antes e depois de 2000

**Fonte:** Elaboração própria

Grau de concordância	Discordo		Nem concordo nem discordo		Concordo	
	Antes de 2000 (%)	Depois de 2000 (%)	Antes de 2000 (%)	Depois de 2000 (%)	Antes de 2000 (%)	Depois de 2000 (%)
Literacia Estatística	0,0	0,0	8,3	8,3	91,7	91,7
Conhecimentos de matemática	8,3	0,0	25,0	16,7	66,7	83,3
Utilizar Sistemas de tratamento de dados	8,3	0,0	0,0	8,3	91,7	91,7
Analisar grandes quantidades de dados	0,0	0,0	25,0	0,0	75,0	100,0
Acompanhar os desenvolvimentos tecnológicos	0,0	0,0	8,3	8,3	91,7	91,7
Cruzar e interpretar dados	0,0	0,0	0,0	0,0	100,0	100,0
Desenvolver visualizações para compreender os dados	0,0	0,0	33,3	8,3	66,7	91,7

Sobre as diferentes capacidades que um jornalista de dados deve desenvolver há um consenso geral, não só entre gerações, mas também entre a literatura e a perspetiva geral dos jornalistas do Expresso. A maioria dos jornalistas do Expresso – de ambos os grupos – concorda que um jornalista de dados deve desenvolver todas as capacidades mencionadas, que se baseiam na literatura. Entre as que reúnem mais consenso para os jornalistas mais e menos experientes está a literacia estatística, a utilização de sistemas de tratamento de dados e o acompanhamento dos desenvolvimentos tecnológicos, os quais reuniram em ambos os grupos cerca de 91,7% de respostas concordantes. E há que destacar uma habilidade específica que todos os profissionais – sem exceção – concordam

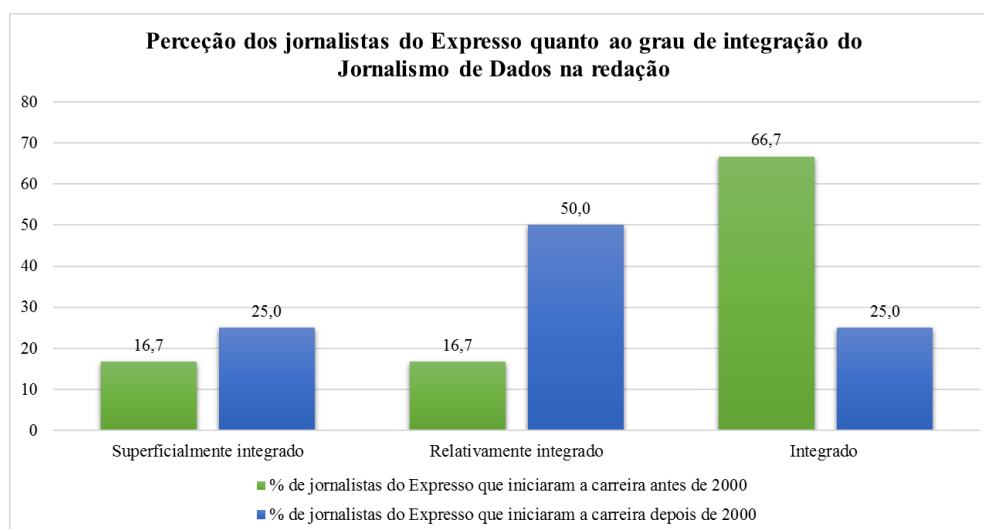


que um jornalista de dados deve desenvolver: cruzar e interpretar dados. Todos os jovens jornalistas (isto é, 100% deste grupo) também concorda que um jornalista deve saber analisar grandes quantidades de dados. No outro grupo, apenas três em cada quatro concorda que esta é uma capacidade que estes profissionais devem desenvolver.

Um outro aspeto mencionado no questionário tem que ver com o grau de integração desta prática jornalística, do jornalismo de dados, na redação do jornal Expresso. Neste sentido, a análise que se segue visa compreender se os jornalistas que somam mais de 18 anos de experiência profissional têm ou não a mesma perspetiva sobre o grau de integração deste fluxo de trabalho na redação que os jornalistas que deram os primeiros passos no jornalismo no presente milénio. No gráfico 5.4, que se segue, estão reunidas as opiniões destas duas gerações de jornalistas do Expresso.

**Gráfico 5.4** – Perceção dos jornalistas do Expresso, que iniciaram a carreira antes e depois do ano 2000, quanto ao grau de integração do jornalismo de dados na redação

Fonte: Elaboração própria



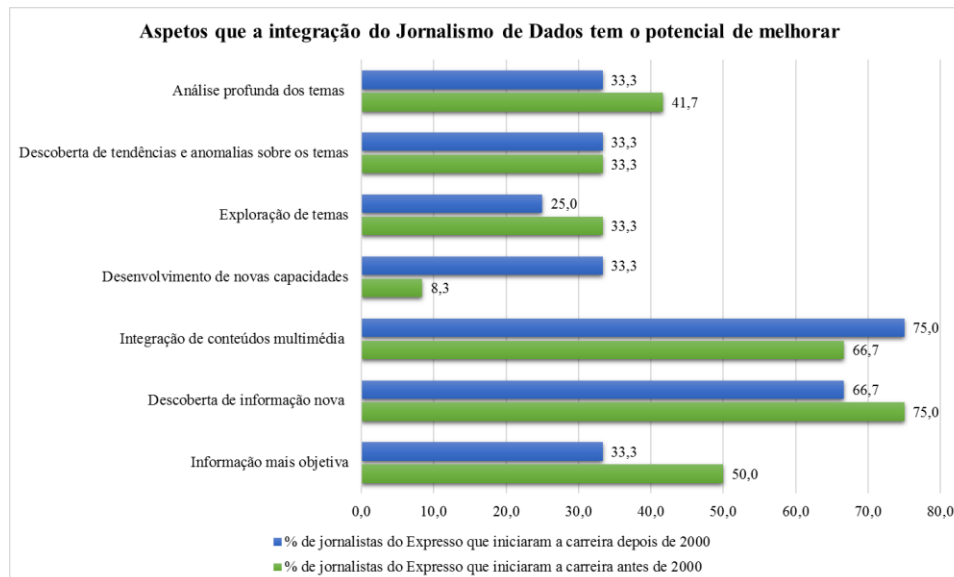
Sobre este aspeto, as opiniões dividem-se. Enquanto a maioria dos jornalistas mais experientes (cerca de 66,7%) considera que o jornalismo de dados está integrado na redação, exatamente metade dos jornalistas mais jovens admite que este está apenas parcialmente introduzido, isto é, num nível intermédio. Apenas um jovem jornalista em cada quatro admite que o jornalismo de dados é parte integrante das rotinas produtivas do jornal Expresso. E, por outro lado, também cerca de 25% deste grupo de jornalistas, dos menos experientes, considera que esta prática é superficialmente adotada. Há menos jornalistas experientes a ter esta visão, de que o jornalismo de dados está superficialmente integrado no seio do jornal Expresso, mas ainda são cerca de 16,7%. Também

aproximadamente 16,7% destes jornalistas, com mais anos de carreira, consideram que o jornalismo de dados está relativamente integrado. As diferentes perceções destes dois grupos de jornalistas deixam a questão inicial em aberto: será o jornalismo de dados uma prática efetivamente integrada no seio da redação do semanário Expresso? Estes resultados indicam apenas que há mais jornalistas experientes a ter a perceção de que o jornalismo de dados é parte integrante da redação do que os jornalistas que iniciaram a carreira depois de 2000.

Um último ponto questionado aos jornalistas do semanário Expresso procura compreender quais são os aspetos que podem ser melhorados – ou que são melhorados – quando o jornalismo de dados é integrado numa redação. Será que as histórias baseadas em dados são mais próximas da objetividade? Será que o jornalismo de dados promove a exploração de temas e a descoberta de tendências e anomalias na perspetiva destes jornalistas? Estas são apenas algumas questões que a análise que se segue se propõe a responder. Os resultados encontram-se compilados no gráfico 5.5.

**Gráfico 5.5** - Aspetos que a integração do jornalismo de dados pode melhorar segundo a perspetiva dos jornalistas que iniciaram a carreira antes e depois de 2000

**Fonte:** Elaboração própria



Embora com algumas nuances – as quais vão ser referidas em seguida –, as opiniões dos dois grupos de jornalistas confluem: a maioria dos jornalistas de ambos os grupos consideram que o jornalismo de dados promove, sobretudo, a descoberta de informação nova e a integração de conteúdos multimédia nas histórias. Os dados mostram também que, para exatamente metade dos jornalistas mais experientes o jornalismo de dados promove a produção de informação mais objetiva e para cerca de 41,7% destes

profissionais a análise profunda dos temas é melhorada. Estes dois últimos aspetos são também nomeados por um jovem jornalista em cada três. A descoberta de tendências e anomalias sobre os temas, através do processo de jornalismo de dados, é um aspeto que reúne especial consenso entre os dois grupos: um em cada três jornalistas considera que este aspeto é melhorado com a integração deste fluxo de trabalho. Também para um terço dos jornalistas mais experientes e para um em cada quatro dos mais jovens, o jornalismo de dados promove a exploração de temas. Enquanto um jovem profissional em cada três admite que adotar esta prática estimula o desenvolvimento de novas capacidades – as quais foram anteriormente referidas –, apenas cerca de 8,3% dos profissionais que mais anos de carreira congregam o referem. Sendo esta uma questão que permitia adicionar outros aspetos, há que destacar uma opinião particularmente construtiva: para um jovem jornalista integrar o processo de jornalismo de dados pode também promover a formação de uma equipa especializada e focada no processo.

Estes são os vários aspetos relativos ao jornalismo de dados sobre os quais é possível caminhar e observar dois lados, duas perspetivas distintas de duas gerações de jornalistas diferentes. Com uns pontos em comum e outros que divergem, as visões dos dois grupos de jornalistas do Expresso serão sintetizadas e confrontadas na discussão que se segue.

#### **5.1.4 Discussão de resultados**

Estudar as perspetivas dos jornalistas do Expresso, dos mais e dos menos experientes, sobre os mais variados aspetos inerentes à prática do jornalismo de dados permitiu perceber, numa primeira instância, que há pontos em que as visões confluem, mas também há outros em que se afastam.

Será que os jornalistas com largos anos de carreira escrevem mais ou menos histórias baseadas em dados do que os jornalistas menos experientes? Os dados mostram uma diferença entre as rotinas dos dois grupos de profissionais: enquanto 66,7% dos jovens jornalistas diz redigir histórias baseadas em dados com frequência ou todos os dias, apenas um em cada quatro jornalistas com largos de carreira assume fazê-lo. Embora os resultados não sejam conclusivos – dada a dimensão da amostra –, revelam um indicador que pode responder à questão supracitada: os jovens jornalistas escrevem notícias baseadas em dados com mais frequência do que os jornalistas mais experientes.

No que diz respeito às dificuldades sentidas pelos jornalistas de ambos os grupos na execução das várias tarefas do jornalismo de dados, verifica-se que tanto há tarefas que possuem um grau de dificuldade semelhante para ambas gerações, como as que possuem graus distintos. Contudo, o número de jornalistas inquiridos e as diferenças apresentadas não são suficientemente expressivas para tirar conclusões. Estas apenas indicam – ou podem indicar – que a análise de dados é a tarefa que apresenta maiores dificuldades para os jornalistas do Expresso de ambos os grupos.

No que diz respeito à origem da informação, os resultados mostram um consenso entre gerações. As fontes mais acedidas pelos jornalistas, que somam pelos menos 18 anos de carreira, são também as mais acedidas pelos jornalistas menos experientes, nomeadamente as organizações governamentais e não-governamentais nacionais e internacionais. Embora com algumas nuances, os principais sistemas de informação utilizados pelos jornalistas do Expresso, de ambas gerações, também são os mesmos: o *Backoffice* e o *Content Station*, usados para gerir e publicar conteúdos nas diferentes plataformas do Expresso. No que diz respeito à análise de dados – um processo crucial na prática em estudo, como já foi referido – o *Microsoft Office Excel* é o sistema mais utilizado, sobretudo pelos jovens jornalistas (75%), mas também pelos mais experientes (41,6%). A maior utilização deste sistema por parte dos jovens jornalistas, em concreto, pode estar de alguma forma relacionada com a frequência de produção de conteúdos baseados em dados, que é maior neste grupo do que no dos jornalistas mais experientes.

As opiniões destas duas gerações de jornalistas do Expresso também convergem quando o assunto diz respeito às capacidades que um jornalista de dados deve desenvolver, nomeadamente: literacia estatística, conhecimentos em matemática, analisar, cruzar e interpretar dados e criar visualizações para os compreender. A maioria dos profissionais do Expresso concordam também que um jornalista de dados deve não só saber utilizar sistemas de tratamento de dados, como também acompanhar os desenvolvimentos tecnológicos que vão surgindo pelo mundo.

A confluência de opiniões deixa de existir quando o assunto é o grau de integração do jornalismo de dados no seio da redação do jornal Expresso. Enquanto a maioria dos jornalistas experientes diz que esta prática está introduzida na redação, três em cada quatro jovens jornalistas consideram que esta está relativa ou superficialmente integrada. Face às diferenças de perceções entre gerações não é possível tirar conclusões que permitam responder à questão inicial, a qual fica, por isso, em aberto – tal como foi anteriormente referido. Posto isto, não é possível perceber, a partir destes resultados, se o jornalismo de dados é ou não uma prática que integra efetivamente a redação do Expresso.

Que aspetos das histórias podem ser beneficiados quando há integração do jornalismo de dados nas redações? Sobre esta questão, há dois pontos que reúnem consenso entre gerações: a descoberta de informação nova e a integração de conteúdos multimédia nas histórias. Para metade dos jornalistas mais experientes, esta prática promove ainda a produção de informação mais objetiva. Entre os aspetos menos nomeados – mas não menos importantes –, encontra-se a exploração e a análise profunda dos temas, a descoberta de tendências e anomalias sobre os mesmos e o desenvolvimento de novas capacidades.

O jornalismo de dados pode significar tudo isto e muito mais. Estes são apenas alguns aspetos – inerentes não só ao processo produtivo, como também ao conteúdo final a ser publicado – que podem ser melhorados. Mas nada melhora sem mudanças, sem percorrer um caminho rumo ao objetivo. Sendo o objetivo melhorar a qualidade dos conteúdos jornalísticos e a confiança do público, o caminho a percorrer é – ou pode ser – explorar as potencialidades do jornalismo de dados. E para tal a redação do Expresso deve contar com o apoio de todos os profissionais de ambas as gerações: com os que possuem mais e menos dificuldades, com os que mais redigem histórias baseadas em dados, mas também com os que escrevem menos.

## **5.2 Qual é a percepção dos leitores do Expresso em relação às histórias baseadas em dados?**

### **5.2.1 Recolha de dados, metodologia e objetivos**

Para investigar as percepções dos leitores portugueses quanto às histórias baseadas em dados foi desenhado um questionário – na aplicação “Formulários do Google” – que foi partilhado com os leitores via *Facebook*, entre 22 de janeiro e 5 de fevereiro de 2018.

O questionário – constituído sobretudo por questões fechadas – foi dividido em três segmentos: um que reúne dados sociodemográficos, outro destinado a conhecer o perfil dos leitores e, por último, um dedicado às histórias baseadas em dados no contexto digital. Tendo por base as 91 respostas obtidas, a investigadora redigiu um artigo, que foi aceite na 13ª Conferência Ibérica de Tecnologias e Sistemas de Informação, a qual ocorreu entre 13 e 16 de junho de 2018 em Cáceres, Espanha. O mesmo – presente no apêndice 5 – foi publicado na biblioteca digital IEEE *Xplore* e enviado para indexação em ISI, *Scopus*, EI, INSPEC e *Google Scholar*.

O objeto de estudo do presente capítulo centra as percepções dos leitores – ou potenciais leitores – do Expresso Digital quanto às histórias baseadas em dados. E, por isso, a investigadora filtrou apenas as respostas dos inquiridos que dizem consultar jornais generalistas no seio digital com frequência ou todos os dias. Partindo do estudo das 45 respostas destes leitores (cerca de 49,5% da amostra) – as quais serão analisadas no *Microsoft Office Excel* e convertidas em visualizações que se encontram no apêndice 4 – será possível perceber as tendências e anomalias nos olhares dos leitores sobre este tema, e, partindo destas pistas, os desafios que o Expresso Digital enfrenta.

### **5.2.2 Dados sociodemográficos**

Tendo em vista uma breve caracterização da seleção de dados acima referida, verifica-se que a maioria dos leitores que consultam com frequência ou todos os dias jornais digitais generalistas (51,1%) – o perfil de leitores referente ao Expresso Digital – têm idades entre os 18 e os 25 anos. E cerca de 48,9% dos leitores possuem idades entre os 26 e os 65 anos. A maioria possui licenciatura (51,1%), cerca de 22,2% mestrado e apenas 2,2% doutoramento. Aproximadamente 17,8% destes leitores completaram o ensino secundário (12ºano) e apenas 6,7% o ensino básico (9ºano). Dentro deste perfil de leitores há ligeiramente mais mulheres (51,1%) do que homens (48,9%).

### **5.2.3 O jornalismo de dados aos olhos dos leitores dos jornais digitais generalistas**

Partindo das respostas dos leitores que dizem consultar com frequência ou todos os dias jornais generalistas no meio digital será possível conhecer não só as suas necessidades, mas também as suas perceções sobre vários aspetos relativos ao jornalismo digital e, em especial, ao jornalismo de dados. Estudar estes resultados significa conhecer as perceções dos leitores – ou potenciais leitores – do Expresso Digital, já que este jornal possui, nada mais nada menos, do que um carácter generalista.

Primeiro, os leitores foram levados a refletir sobre os vários aspetos referentes ao jornalismo digital que podem ser melhorados (apêndice 4.1). Entre os vários pontos referidos, há cinco que possuem especial consenso entre os leitores assíduos de jornais digitais generalistas. A grande maioria destes inquiridos concorda que se deve investir em novas formas de interação com o conteúdo (86,7%), adaptar o conteúdo a diferentes plataformas, como *smartphones*, *ipads* (86,7%), apostar na contextualização dos temas (84,4%), inserir mais conteúdos multimédia (77,8%), e consideram também que é importante envolver mais os leitores no jornalismo recolhendo as suas opiniões (55,6%).

Com que frequência estes leitores consultam histórias baseadas em dados? A maioria destes inquiridos (57,8%) diz ler estas notícias com frequência ou todos os dias, cerca de 24,4% admite fazê-lo apenas por vezes e apenas 17,8% raramente ou nunca. Os resultados (reunidos no apêndice 4.2) indicam – ou podem indicar – que a maioria dos leitores, com o perfil em estudo, consulta com regularidade notícias baseadas em dados.

Perante várias afirmações sobre o jornalismo de dados, os leitores foram levados a refletir sobre os aspetos que mais e menos concordam (apêndice 4.3). Para a generalidade dos leitores assíduos de jornais digitais generalistas (88,9%), os dados tornam as notícias mais credíveis. A grande maioria (84,4%) concorda que as histórias baseadas em dados melhoram a sua perceção da realidade, pois é possível quantificar ou qualificá-la, e admitem também explorar o conteúdo para compreender melhor o significado dos dados. Para três inquiridos em cada quatro, os dados tornam a informação mais próxima da objetividade. Mais de metade (51,1%) destes inquiridos concorda que geralmente estes conteúdos possuem mais visualizações. Oito leitores em cada dez admitem reconhecer que as histórias baseadas em dados resultam de um trabalho diferenciado e, por isso, atribuem-lhe valor. E sete em cada dez consentem que estas notícias têm qualidade acrescida. A maioria destes leitores (64,4%) admite ainda confiar mais na informação quando contem dados.

Foram várias as afirmações sobre as histórias baseadas em dados – construídas partindo da literatura – que fazem sentido para os leitores de jornais digitais generalistas. Mas quais são, afinal, os aspetos que podem ser melhorados no meio digital quando se investe – ou reforça uma aposta – em jornalismo de dados? A maioria destes leitores nomeia, especialmente, três aspetos (apêndice 4.4). Mais de metade (51,1%) considera que a visualização de conteúdos é aprimorada. Cerca de 73,3% concorda que a objetividade da informação é melhorada. E, por fim, a grande maioria destes leitores (77,8%) assume que esta pode ser uma via para melhorar a confiança dos leitores no jornalismo. Sendo esta uma questão que acolhe sugestões, foram registados mais dois aspetos pelos inquiridos. Um deles assume que investir em jornalismo de dados promove a solução de problemas. E outro diz que melhora a assertividade jornalística.

#### **5.2.4 Um caminho possível: reforçar a aposta em jornalismo de dados**

Cruzando as necessidades dos leitores regulares de jornais digitais generalistas com as suas perceções verifica-se que há – ou pode haver – um caminho a percorrer. No caso do Expresso Digital, reforçar a aposta já iniciada em jornalismo de dados – através da rubrica 2:59, por exemplo – pode colmatar algumas necessidades apresentadas pelos seus leitores ou potenciais leitores. Construir mais histórias baseadas em dados, incorporando visualizações interativas, poderá, por exemplo, colmatar duas necessidades apontadas: incorporar mais conteúdos multimédia e melhorar a interação dos leitores com a história, um aspeto também indicado na literatura (nomeadamente no subcapítulo 2.3).

Acima de tudo, estes resultados revelaram que, na ótica deste perfil de leitores, o jornalismo de dados pode ter o poder de melhorar vários aspetos deteriorados pela crise que hoje paira sobre o jornalismo, os jornalistas e os seus leitores. Para a maioria destes leitores, as histórias baseadas em dados têm o poder de melhorar a exatidão e a credibilidade da informação, aprimorando a sua qualidade. E têm também o poder de melhorar a sua perceção da realidade. Todos estes aspetos positivos podem, quem sabe, abrir caminho para reanimar a identidade jornalística. E, ao mesmo tempo, podem também ajudar a reverter o clima de desconfiança existente na profissão. O jornalismo de dados tem – ou pode ter – o poder de melhorar a confiança do público no jornalismo.





## CONCLUSÃO

As novas tecnologias e a rede que hoje interliga o mundo enraizaram-se na esfera do jornalismo, transformando-o por completo. É neste contexto que o jornalismo digital ganha vida e abraça uma série de possibilidades. E também é nele que o jornalismo procura novas formas de assegurar a manutenção da sua credibilidade. A expansão deste universo digital cria um enorme desafio para a profissão: compreender os bilhões de dados que todos os dias são gerados. E este não só desencadeia a evolução da prática jornalística que mergulha nos dados para trazer novas histórias, como também evidencia sua importância na atualidade.

Num ambiente marcado pela instabilidade, os *media* viram no jornalismo de dados uma oportunidade, por várias razões. Hoje, as organizações mediáticas têm o acesso facilitado à principal matéria-prima, aos dados, que são disponibilizados *online*, periodicamente, por entidades oficiais. E têm também acesso às ferramentas para analisar e interpretá-los, às mais recentes tecnologias e sistemas de informação. Acima de tudo, os *media* viram neste fluxo de trabalho bem definido – em que há recolha, tratamento, análise, interpretação e visualização de dados – o caminho para desmistificar o mundo social que está escondido nos dados, desvendando, assim, histórias novas com valor agregado.

Da sua aposta em jornalismo de dados nascem novos projetos digitais: o diário britânico *The Guardian* investe no *DataBlog*, o *ProPublica* abraça a *DataStore* e o jornal Expresso dá vida à rubrica 2:59. E, com a integração destes projetos, surgem também uma série de mudanças na estrutura das redações, nas rotinas dos jornalistas e na percepção dos leitores. Nas redações, assiste-se à materialização do modelo de negócio delineado pela direção. Assiste-se, portanto, à reutilização de recursos tecnológicos e à introdução de outros. Assiste-se também à atribuição de novas funções aos profissionais, mas também ao recrutamento de outros. Assiste-se, assim, à criação de uma equipa multidisciplinar que se encarrega de assegurar todo o processo de construção do projeto, o qual se propõe a descobrir os laços existentes entre um mar de dados. Os jornalistas, que mergulham no universo dos dados, passam a incluir outras tarefas nas suas rotinas, como é o caso da recolha, tratamento, análise e interpretação de dados. E, por esse motivo, devem cultivar capacidades específicas, entre as quais: conhecer a natureza dos dados, desenvolver literacia estatística e, sobretudo, saber manusear os sistemas de tratamento e manipulação de dados. O público, que interage com a história final, tem a oportunidade de compreender um tema complexo em profundidade, navegando entre os dados segundo uma perspetiva

pessoal, o que estimula – ou pode estimular – a criação de uma empatia com a história. As possibilidades que o jornalismo de dados traz para os leitores são tantas, que este já é considerado um caminho possível para melhorar a sua confiança no jornalismo.

Parte das mudanças mencionadas foram percecionadas no seio da redação do jornal Expresso, no decorrer dos seis meses de estágio. E ainda foram reconhecidas outras alterações particulares, sobretudo, na realização da rubrica 2:59. Quase todas – senão todas – as fases do processo decorrem na redação, as quais são suportadas não só pelos recursos tecnológicos já existentes (como é o caso do *Excel* para analisar dados), mas também por novas ferramentas, como é exemplo o *Tableau Software*, utilizado para desenvolver visualizações interativas. A maioria dos jornalistas do Expresso estão – ou devem estar – aptos intelectualmente para aplicar o processo de jornalismo de dados, dado que podem surgir bases de dados novas sobre temas que se inserem nas respetivas áreas de especialização. E, quando os prazos permitem, os jornalistas trabalham em conjunto com infografistas para enriquecer a história com visualizações baseadas em dados. O carácter multidisciplinar inerente ao processo é ainda mais evidente na construção da rubrica 2:59. Para além de editores, jornalistas e infografistas, o 2:59 conta também com coordenadores, realizadores, operadores de câmara, *designers*, entre outros profissionais. Esta rubrica reflete também a importância da convergência mediática no seio digital, visto que é nos estúdios da SIC que a equipa grava, todas as semanas, o 2:59.

São vários os atores que interferem no processo, sendo que cada um exerce funções próprias e importantes nas suas várias fases. Mas são os jornalistas que possuem um papel central. São eles que recolhem e analisam os dados, e desenham o rumo de uma história. Sobre o que muda nas rotinas e na vida destes profissionais quando abraçam o jornalismo de dados, os resultados obtidos através do questionário direccionado aos jornalistas do Expresso revelam indicadores importantes, mas não permitem retirar conclusões definitivas, dado o pequeno número de respostas obtidas – a taxa efetiva de resposta foi inferior a 50%. Estes resultados indicam que a adaptação a este processo rigoroso apresenta dificuldades para os jornalistas, sobretudo na fase de análise de dados. E revelam também que estas novas tarefas requerem o desenvolvimento de diferentes saberes, como aplicar técnicas e métodos de análise de dados e manusear sistemas de informação que suportam o seu tratamento. Estes resultados indicam ainda que o jornalismo de dados é – ou pode ser – um novo caminho possível para descobrir informações novas, que de outro modo não seriam desvendadas.

Os aspetos que mudam – ou que podem mudar – para os leitores do Expresso quando leem histórias baseadas em dados também foram reunidos e estudados, embora com limitações. Tendo em conta o tamanho do universo em estudo – todos os leitores de jornais generalistas *online* –, a amostra (n=45) torna-se pequena para tirar conclusões. Ainda assim, estes dados revelam tendências que podem indicar vontades. Os resultados indicam, genericamente, que os dados transmitem confiança aos leitores e tornam as histórias mais rigorosas, credíveis e próximas da realidade. E revelam também que os leitores tendem a explorar as histórias para melhor compreender o significado dos dados.

Para além das várias pistas referidas, que alertam para as possíveis mudanças associadas à integração do jornalismo de dados, os seis meses de estágio no seio do jornal Expresso revelaram outras, que mostram a importância dos dados – gerados em ambiente digital – na construção de estratégias de negócio. Esses dados, registado no *Google Analytics*, mostram tendências de leitura, evidenciando os temas mais e menos lidos. E partindo do pressuposto que há uma relação direta entre visualizações e receitas no seio digital, as decisões tomadas a nível organizacional procuram dar destaque aos temas que mais atraem o público para, assim, otimizar as receitas obtidas. No Expresso, esta lógica está presente ao nível das estratégias de *marketing* digital adotadas: as notícias republicadas ou destacadas nas redes sociais são as que mais reações e visualizações geram. E também são detetadas influências da referida lógica nas decisões tomadas em matéria editorial. Este ponto é claro no processo de análise e seleção de temas que merecem ser desenvolvidos segundo novas perspetivas: os eleitos são os que registam mais leituras no *Google Analytics* e, portanto, os que mais interesse despertam no público. As decisões a nível editorial assentes nas tendências de leitura podem indicar, por outro lado, que o Expresso possui uma preocupação em adequar os seus conteúdos às novas necessidades e interesses dos seus leitores. Ainda assim, na perspetiva da investigadora, estes dados podem ser um ponto de partida para uma reflexão maior, porque revelam que há temas importantes que não estão a ser lidos. Partindo destas pistas, os editores podem, por exemplo, refletir sobre os motivos que ditam o insucesso das histórias (o título pouco claro, falta de contextualização da entrada), para depois melhorá-las e promovê-las.

O futuro sinaliza caminhos possíveis para dar um novo fôlego ao jornalismo, que hoje se encontra em crise. A necessidade de compreender o mar de dados hoje disponível no meio digital, leva o jornalismo mais longe, abrindo uma janela de oportunidade. Leva-o tão longe que o pode ajudar a recuperar a sua identidade, marcada pelo rigor, pela verdade e

pela incansável vontade de procurar factos que transmitem certezas. E pode também ajudar a melhorar a relação de confiança existente entre os leitores e o jornalismo, que de momento se encontra danificada, melhorando a sua perceção da realidade e da credibilidade da informação. O jornalismo de dados tem o poder de revelar o mundo complexo escondido nos dados. Mas pode ter um poder ainda maior, pode dar um novo ânimo ao jornalismo, ajudando-o a reverter a sucessão de crises que hoje pairam sobre a profissão. Primeiro, auxiliando a restaurar a identidade jornalística. Depois, ajudando a melhorar a confiança do público no jornalismo. E, por conseguinte, apoiando o desenvolvimento da estabilidade económica nas redações.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Agranonik, M., & Hirakata, V. N. (2011). Sample size calculation: proportions. *Revista HCPA*, 31(1), 382–388. Retrieved from <http://seer.ufrgs.br/hcpa>
- APCT. (2018). Analise Simples: a Circulação Total Impressa que inclui a Circulação Total Digital. Retrieved August 7, 2018, from [http://www.apct.pt/Analise\\_simples.php](http://www.apct.pt/Analise_simples.php)
- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden: Introducing new methods and genres of journalism into “old” organizations. *Digital Journalism*, 2(3), 394–405. <https://doi.org/10.1080/21670811.2014.884344>
- Baak, S. (2011). A new style of news reporting : Wikileaks and Data-Driven Journalism. *SSOAR*.
- Bastos, H. (2007). Da implementação à estagnação : os primeiros doze anos de ciberjornalismo em Portugal. In *Bocc.ubi*.
- Bounegru, L. (2012). Data Journalism in Perspective. Retrieved July 6, 2018, from [http://datajournalismhandbook.org/pt/introducao\\_4.html](http://datajournalismhandbook.org/pt/introducao_4.html)
- Bradshaw, P. (2011a). 6 ways of communicating data journalism. Retrieved July 24, 2018, from <https://onlinejournalismblog.com/2011/07/13/the-inverted-pyramid-of-data-journalism-part-2-6-ways-of-communicating-data-journalism/>
- Bradshaw, P. (2011b). The Inverted Pyramid of Data Journalism. Retrieved July 18, 2018, from <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>
- Bradshaw, P. (2012). What Is Data Journalism? Retrieved July 18, 2018, from [http://datajournalismhandbook.org/1.0/en/introduction\\_0.html](http://datajournalismhandbook.org/1.0/en/introduction_0.html)
- Canavilhas, J. (2003). Webjornalismo. Considerações gerais sobre jornalismo na web. In *BOCC*. Retrieved from [https://ubibliorum.ubi.pt/bitstream/10400.6/4358/1/CAPÍTULO\\_WebjornalismoConsideraçõesgerais.pdf](https://ubibliorum.ubi.pt/bitstream/10400.6/4358/1/CAPÍTULO_WebjornalismoConsideraçõesgerais.pdf)
- Canavilhas, J. (2005). Os jornalistas online em Portugal, 1–8.
- Canavilhas, J. (2006). Webjornalismo: Da pirâmide invertida à pirâmide deitada. In *BOCC*. Retrieved from <http://www.bocc.ubi.pt/~boccmirror/pag/canavilhas-joao->

webjornalismo-piramide-invertida.pdf

- Coddington, M. (2015). Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331–348. <https://doi.org/10.1080/21670811.2014.976400>
- Coelho, P. M. R. S. (2013). *A formação académica para o jornalismo do século XXI: sobre questões de prática e técnica Jornalismo e mercado - os novos desafios colocados à formação*. Retrieved from <http://hdl.handle.net/10362/12109>
- Cohen, S. (2012). Visualization as the Workhorse of Data Journalism. Retrieved July 24, 2018, from [http://datajournalismhandbook.org/1.0/en/delivering\\_data\\_3.html](http://datajournalismhandbook.org/1.0/en/delivering_data_3.html)
- Cohen, S., Li, C., Yang, J., & Yu, C. (2011). Computational Journalism: a call to arms to database researchers. *Proceedings of the 5th Biennial Conference on Innovative Data Systems Research (CIDR 2011) Asilomar, California, USA.*, (January), 148–151. Retrieved from [http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/en//pubs/archive/37381.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en//pubs/archive/37381.pdf)
- Corrêa, E. S., & Lima, M. C. (2013). Negócios de mídia na era da midiatização: uma reflexão sobre modelos de exposição e interação. In H. Bastos & F. Zamith (Eds.), *Ciberjornalismo - Modelos de negócio e redes sociais*. Edições Afrontamento.
- Correia, F. (2006). *Jornalismo, Grupos Económicos e Democracia*. Caminho.
- Davis, C. (2012). How to Build a News App. Retrieved July 26, 2018, from [http://datajournalismhandbook.org/1.0/en/delivering\\_data\\_1.html](http://datajournalismhandbook.org/1.0/en/delivering_data_1.html)
- Expresso. (2016). Gafes, erros, quedas e até danças: como se faz o 2:59. Retrieved August 7, 2018, from <https://expresso.sapo.pt/multimedia/259/2016-07-28-Gafes-erros-quedas-e-ate-dancas-como-se-faz-o-2#gs.bTlu66M>
- Expresso. (2018). As fotolegendas que foram uma revolução. *Expresso Diário*. Retrieved from <https://leitor.expresso.pt/diario/05-01-2018/html/caderno-1/temas-principais/as-fotolegendas-que-foram-uma-revolucao-1-1>
- Expresso, & Sofia, R. (2017). Mulheres a mandar nas autarquias, procuram-se. Em 1976 eram cinco, hoje são 31. *Expresso*. Retrieved from <https://expresso.sapo.pt/autarquicas2017/2017-10-17-Mulheres-a-mandar-nas->

autarquias-procuram-se.-Em-1976-eram-cinco-hoje-sao-31#gs.B8K=95k

- Fernandes, J. M. (2000). A Génese da Máquina da Sociedade de Informação-Baby Machine, o protótipo dos actuais computadores. Évora/ Aveiro: 1º Congresso Luso-Brasileiro de Ciência e da Técnica. Retrieved from <http://www.cs.iastate.edu/jva/jva-archive.shtml>
- Figueira, J. (2012). *O Essencial Sobre a Imprensa Portuguesa (1974-2010)*. Angelus Novus.
- Garcia, J. L. (2009). Introdução ao estudo dos jornalistas portugueses - Os jornalistas e as contradições do capitalismo jornalístico no limiar do século XXI. In *Estudo sobre os jornalistas portugueses: Metamorfoses e encruziladas no limiar do século XXI* (pp. 23–46). Imprensa de Ciências Sociais.
- Geere, D. (2012). Engaging People Around Your Data. Retrieved August 6, 2018, from [http://datajournalismhandbook.org/1.0/en/delivering\\_data\\_10.html](http://datajournalismhandbook.org/1.0/en/delivering_data_10.html)
- Gomes, R. M. (2008). O papel da Internet no contacto entre jornalistas e fontes de informação nas redações portuguesas. *Revista Do Centro de Investigação Media e Jornalismo - Um Século de Ensino Do Jornalismo*, 37–53.
- Google. (2018). Google News Lab. Retrieved August 5, 2018, from <https://newsinitiative.withgoogle.com/google-news-lab>
- Graça Meireles, S. (2007). Uma Introdução ao Estudo dos Jornalistas Portugueses. In *Os Jornalistas Portugueses: Dos problemas da inserção aos novos dilemas profissionais* (pp. 17–28).
- Henninger, M. (2013). Data-Driven Journalism. In *Challenge and Change. Reassessing Journalism's Global Future* (pp. 157–184).
- Howard, A. (2014). The Art and Science of Data-Driven Journalism. *Tow Centre for Digital Journalism - a Tow/Knight Report*. <https://doi.org/10.1017/CBO9781107415324.004>
- Kayser-Bril, N. (2012). Become Data Literate in 3 Simple Steps. Retrieved July 28, 2018, from [http://datajournalismhandbook.org/1.0/en/understanding\\_data\\_0.html](http://datajournalismhandbook.org/1.0/en/understanding_data_0.html)
- Kirkpatrick, K. (2015). Putting the Data Science into Journalism. *COMMUNICATIONS OF THE ACM*, 58(15). <https://doi.org/10.1145/2742484>



- Klein, S. (2012). News Apps at ProPublica. Retrieved July 26, 2018, from [http://datajournalismhandbook.org/1.0/en/delivering\\_data\\_2.html](http://datajournalismhandbook.org/1.0/en/delivering_data_2.html)
- Lewis, S. C. (2015). Journalism In An Era Of Big Data. *Digital Journalism*, 0811, 321–330. <https://doi.org/10.1080/21670811.2014.976399>
- López-García, X., Toural, C., & Rodríguez-Vázquez, A.-I. (2016). Software, estadística y gestión de bases de datos en el perfil del periodista de datos. *El Profesional de La Información*, 25(2), 286. <https://doi.org/10.3145/epi.2016.mar.16>
- Lorenz, M. (2012). Business Models for Data Journalism. Retrieved July 18, 2018, from [http://datajournalismhandbook.org/1.0/en/in\\_the\\_newsroom\\_10.html](http://datajournalismhandbook.org/1.0/en/in_the_newsroom_10.html)
- Mogoş, A. (2016). Data Analysis and Visualisation. Theory and Tools for Data Journalists. In C. Singla, I. Da Rocha, & X. Ramon (Eds.), *Shaping The Future of News Media* (pp. 59–92). Integrated Journalism in Europe European Commission.
- Moherdaui, L. (2017). *Jornalismo sem manchete: a implosão da página estática*. São Paulo: Editora Senac São Paulo. Retrieved from <https://www.amazon.com.br/Jornalismo-sem-manchete-implosão-estática-ebook/dp/B074FZV2MV>
- Oliveira, A. C. P. de. (2012). *O Jornal Expresso e o Processo de Convergência dos Media: das Plataformas aos Conteúdos*. Universidade Nova de Lisboa.
- Pavlik, J.(2014). The News Code: Implications of Data, Algorithms and Connectivity for Journalism Quality in the Digital Age. In *Congresso Internacional de Ciberjornalismo* (pp. 8–20). Porto: Observatório do Ciberjornalismo.
- Picard, R. (2013). *Criação de Valor e o Futuro das Organizações Jornalísticas: porque motivo e como o Jornalismo deve mudar para se manter relevante no século XXI*.
- Pissarra Esteves, J. (2007). Internet e Democracia: Estados Sociedade Civil perante os novos desafios da comunicação política. *Revista Do Centro de Investigação Media e Jornalismo - Jornalismo e Actos Da Democracia*, 33–42.
- Pombo, C., Sousa, V., Atayde, A., Beleza, J., & Paes, C. (2018). 2:59 - Já bebeu café hoje? *Expresso*. Retrieved from <https://expresso.sapo.pt/multimedia/259/2018-04-26-Ja-bebeu-cafe-hoje-#gs.LJmHxU0>
- Prodanov, C. C., & Freitas, E. C. de. (2013). *Metodologia do Trabalho Científico* :

- Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico*. Universidade Feevale.
- Ramos, D., Spinelli, E., & Estevanim, M. (2015). Formatos Digitais e Qualidade no jornalismo Brasileiro em Base de Dados. In *Iv International Conference on Online Journalism* (pp. 36–49).
- Rebelo, J. (2014). *As Novas Gerações de Jornalistas em Portugal*. Lisboa: Editora Mundos Sociais.
- Rodrigues, I. B. (2015). *A Era do Jornalismo Digital: o Caso do Jornal Expresso que ao semanário juntou uma edição diária*. Universidade Católica Portuguesa.
- Rosenbaum, M. (2012). Data Stories. In J. Gray, L. Chambers, & L. Bounegru (Eds.), *The Data Journalism Handbook*. O'Reilly Media.
- Cohen, S. (2012). Using visualizations to Tell Stories. Retrieved July 24, 2018, from [http://datajournalismhandbook.org/1.0/en/delivering\\_data\\_4.html](http://datajournalismhandbook.org/1.0/en/delivering_data_4.html)
- Souza, R. M. (2017). Investigando as fake news: análise das as agências fiscalizadoras de notícias falsas no Brasil. In *XXII Congresso de Ciências da Comunicação na Região Sudeste*. Intercom – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação.
- Teixeira, W., & Junior, L. (2012). Big Data, Jornalismo Computacional e Data Journalism: estrutura, pensamento e prática profissional na Web de dados. *Estudo Em Comunicação* Nº 12, 207–222.
- Traquina, N. (2007). A identidade jornalística. In *O que é Jornalismo* (2ª edição, pp. 130–161). Quimera.
- Trasel, M. (2014). Aprendendo a se deixar guiar por dados: a formação dos jornalistas da equipe Estadão de Dados. *Revista Brasileira de Ensino de Jornalismo*, 86–99.



## **APÊNDICES**

## **APÊNDICE 1. Entrevista ao editor adjunto do Expresso Diário José Cardoso**

Entrevista realizada no dia 6 de dezembro 2017

### **1. Como se processa a seleção de temas para o Diário e para o site?**

Para além das propostas que os vários jornalistas fazem, baseamo-nos também nas agendas: na agenda da lusa, que tem as previsões para os dias seguintes. E vemos o que é que vai acontecer, o que é previsível. Baseamo-nos também na informação que recolhemos da leitura dos jornais do dia, dos jornais da manhã. As sugestões ou as ideias para os temas do dia aparecem de todos os lados e de todas as maneiras, mas é essencialmente por propostas das pessoas, dos jornalistas, das agendas, essencialmente a da lusa, e leitura dos jornais do dia ou da véspera.

### **2. Não há uma reunião matinal onde selecionam os temas que vão para o Diário?**

Para o Diário sim, embora seja um processo muito plástico. O site está sempre a ser alimentado, essencialmente, pelas pessoas que trabalham para o *site* mas também pelo restantes jornalistas que acompanham os temas sobre os quais se está a falar ou sobre quais vai havendo notícias. O *site* está sempre a ser alimentado em contínuo. Já o diário não. Nós decidimos de manhã, muitas vezes decidimos na véspera, e se soubermos que algo vai acontecer daqui a três dias, colocamos logo na nossa própria agenda. Pode-se ir preparando. Para o Diário é essencialmente no dia. Mas é uma coisa muito plástica no sentido em que há dias em que temos logo de manhã sete ou oito temas principais para o Diário, mas há outros em que, de manhã, só temos um ou dois e temos que andar a ver o que pode haver de notícia nesse dia. E podemos chegar o fim da tarde, mais ou menos na altura de saída do Diário, e temos à mesma sete temas, mas são apenas dois dos que estavam previstos de manhã, porque, entretanto, surgiram outras histórias, e vamos substituindo. E no caso de algumas notícias que nós tínhamos para o Diário, que sai às 18 horas, serem ‘queimadas’ por outros sites, publicamo-las logo no nosso *site*. O que acontece muitas vezes – ou todos os dias – é que pegamos no noticiário da agenda e escrevemos uma notícia para o *site*, e depois vamos desenvolvê-la de outra maneira, com alguma mais valia, para o Diário. Por exemplo, neste momento o banco de Portugal está a apresentar um relatório sobre economia nacional que é importante. Temos lá uma jornalista, ela quando vier, vai escrever já uma notícia para o *site*, com os principais

indicadores, e depois vai desdobrar o relatório todo para ver o que é que pode extrair dali, para ver se há alguma informação interessante que esteja lá escondida. Caso a encontre e desenvolve-a para o Diário. E, portanto, o leitor do Diário tem uma mais valia, porque tem mais informação do que os *sites*, nomeadamente o nosso, lhe dão.

### **3. Quais são as necessidades dos leitores que o Expresso Diário visa preencher atualmente?**

É a questão do papel dos semanários, que com o correr do tempo têm vindo a perder leitores, protagonismo, importância e com o fluxo noticioso permanente é muito difícil para um semanário ter notícias em primeira mão e boas notícias, trabalhos bem desenvolvidos e muitos para encher um semanário. E há muitas notícias que não aguentam ao contrário do que acontecia há 20 anos, em que sabíamos, se a nossa fonte fosse segura, que até sábado ninguém ia dar a notícia. Agora é muito difícil. Muitas vezes passados dez minutos já alguém está a dar. O ritmo noticioso não se compadece com um semanário apenas e, portanto, o que nós fizemos foi criar um jornal diário, que é digital, que dá a informação de um diário, ou seja, tem notícias novas, todos os dias. E sai ao fim da tarde, 18 horas, porque é quando as pessoas têm mais disponibilidade. Não quer dizer que o leiam logo as 18 horas, até porque há um pico depois grande por volta das 21h, 22 horas da noite. A ideia é darmos um bocado ao leitor, que é assinante do Expresso, o diário do dia seguinte, isto porque vamos ver depois os jornais diários no dia seguinte – como o Diário de Notícias, do Público e do jornal I – e muitos dos temas que lá vêm são os temas que nós tratámos. Isto acontece porque temos bons jornalistas, boa informação e boas fontes. Eles têm um processo de produção mais longo, em papel, ou seja, têm que fechar para enviar o jornal para a gráfica e para depois ser distribuído pelas bancas. O Diário é imediato porque é digital. E isso também contribui muito para que nós tenhamos informação exclusiva e em primeira mão às 18h da tarde de um dia.

### **4. Como é que os temas são distribuídos?**

São distribuídos de várias maneiras. Primordialmente para quem trata habitualmente desse tema, portanto tem *know-how* para escrever sobre. Como isso muitas vezes não é possível, por vários motivos, é entregue aos jornalistas que têm mais disponibilidade, a quem está de ‘piquete’, a quem pode. Por exemplo, eu hoje queria um tema sobre a mudança da embaixada dos Estados Unidos de Telavive para Jerusalém, uma das pessoas mais habilitadas para escrever sobre isso seria a Margarida Mota, que escreve sobre

internacional e sobre o Médio Oriente. Contactei-a logo de manhã, mas está de folga hoje, não pode. Portanto falei com alguém do internacional, que esteja dentro do assunto. Mas como o internacional está hoje a fechar a edição em papel, os seus jornalistas também não têm disponibilidade. Também não posso estar a ‘roubar’ os jornalistas todos para trabalharem no Diário de hoje e descorar a edição em papel do sábado. Portanto pedi a uma pessoa que escreve habitualmente no internacional, mas que não está aqui na redação, que é a Joana Azevedo Viana. Depende muito, depende de quem pode. Portanto, em primeira mão atribuem-se os temas a quem geralmente trata dos assuntos. Se for educação para quem trata de educação, se for saúde para quem trata de saúde. Mas como isso nem sempre é possível - e muitas vezes não é - é para quem pode. E quem tem condições para.

**5. Quando o diário foi lançado em 2014, houve por exemplo contratações de jornalistas para sustentar o diário?**

Zero. O Diário é feito com as pessoas da redação só.

**6. E as pessoas da redação são versáteis para escrever para qualquer formato?**

Sim, há pessoas que escrevem em determinadas situações, quando há picos noticiosos sobre determinados assuntos que elas cobrem habitualmente. Escreve duas ou três coisas para o *site*, uma para o Diário e escrevem também para o semanário. A grande dificuldade é a distribuição do noticiário: o que é que vai para o papel, o que vai para o Diário. Essa é a principal luta que nos temos diariamente. Às vezes arriscamos e, às vezes falhamos.

**7. Considera que a escrita deve ser adaptada ao meio digital?**

Muitas vezes já o fazemos, depende da história. Depende de várias coisas: primeiro depende da história, porque se for uma notícia pura e dura, como é o caso do tema sobre a conspiração para matar a primeira ministra britânica Theresa May, que sairá hoje no Diário, não podemos começar esse texto com a pirâmide invertida. Não vamos dizer um tipo foi levado a tribunal hoje acusado de conspiração islâmica para matar Theresa May, porque isso estará horas antes no *site*. Portanto não podemos começar assim, temos que arranjar uma forma de escrever isto de outra maneira. O foco vai ser diferente. Ou investigamos mais alguma coisa, ou tentamos escrever isto como se fosse um romance policial, como fazemos muitas vezes. Depois aqui também entra a qualidade da escrita, para conseguirmos agarrar o leitor com uma história que seja construída de outra maneira.

Muitas vezes não se deve respeitar o princípio da pirâmide invertida, no caso do Diário. Porque muitos desse noticiário já está ‘queimado’ nos sites.

**8. Um texto longo é facilmente lido pelas gerações mais novas, ou depende da história?**

Depende da história. No essencial é uma luta que eu tenho diariamente com muita gente é dar-lhes na cabeça para não escreverem tanto. Porque ou é uma boa história e muito bem escrita, contada, muito bem ilustrada, seja com imagem seja com vídeo, e aí pode ter leitura. Se for um texto, como nós publicamos muitas vezes até no site, com 8 ou 9 mil caracteres corrido, com uns subtítulos e tal, eu acho que o pessoal chega a meio e desiste, sobretudo, a malta mais nova. Não há hipótese, por muito que as pessoas digam que querem informação e quanto mais informação têm mais querem, eu não acredito nisso. Ou seja, os textos devem ser mais curtos.

**9. Que alterações podem ser feitas para atrair mais leitores?**

Investir em multimédia. A informação pode até nem ter texto nenhum. Dá-se o exemplo do 2:59 para explicar o mundo. Pode não ter informação escrita praticamente nenhuma, tem é de ter informação: seja em fotos seja em vídeo, seja em hiperligações. Nós usamos pouco. Usamos de menos, devíamos usar a mais, por vários motivos. No *site*, por alguma falta de cultura, isto é, falta de prática das pessoas, que ainda estão muito formatadas para o papel ou não estão muito formatadas para recorrerem a tudo aquilo que há ao nosso alcance para dar a informação de outro modo sem ser em prosa. No Diário também, essencialmente por questões técnicas: até há pouco tempo não conseguíamos por um vídeo do *youtube*. Nós não conseguimos fazer uma imagem em movimento. Já temos experimentado várias vezes, nomeadamente na capa do Diário. Consegue-se por tabelas e mapas interativos, mas há também limitações técnicas, e a nível da concretização pela secção de infografia. Se pedir um mapa interativo às 8h da manhã para às 16h, por exemplo, se calhar já não há tempo, pois demora muito tempo a fazer. Talvez a solução seja um mapa estático. Devíamos ter muito mais, quer no site quer no Diário.

**10. Como considera o grau de interação do jornal Expresso com os seus leitores?  
O que pode ser melhorado?**

Dizem que podia ser muito melhorado.



**11. Porque é que o site não tem, por exemplo, um espaço para comentários?**

Já teve, mas acabou-se, porque questões praticas, ou seja, havia fluxo permanente de comentários e muitos deles, eram comentários obscenos. Enquanto houve capacidade para moderar tivemos. Havia uma pessoa responsável por ler e apagar os ofensivos e responder. Depois houve uma altura que deixou de ser possível, porque as pessoas cobertas de anonimato dizem tudo e mais alguma coisa e, por isso, teve de se acabar com isso porque senão era uma obscenidade pegada. Fazem criticas destrutivas e dizem asneiras. Já não havia capacidade de moderação no meio de centenas de comentários diários e os leitores do Expresso também não podiam ficar à mercê de insultos, comentários sem pés nem cabeça E, por tudo isso, acabou-se com os comentários. Agora era importante, era.

**12. Então, o leitor não tem resposta?**

Não, não há.

**E não seria importante?**

E importante, mas como é que isso se faz sem contratar mais pessoas? A interação com os leitores dá-se através da leitura e também através de comentários e de *email* que os leitores enviam para os jornalistas e para os editores. Cada um tem responsabilidade de responder. Há ali uma interação. Muitas vezes há leitores que fazem determinada crítica negativa e depois veem que não tinham razão nenhuma. Porque leem o *lead* mas não leem o resto e mandam um *email* a descascar. A voragem da informação é tal que as pessoas já não leem tudo até ao fim, e depois fazem comentários tento em conta apenas uma rama.

**13. Considera que uma maior aposta na interação com os leitores poderia aproximar o jornal dos mesmos?**

Sim aproximaria certamente. O problema é quem gere os orçamentos e onde decidem investir.

**14. Porque é que o Diário está intrinsecamente ligado ao Semanário – através do voucher?**

O facto de existir o Diário e de ser necessário ou uma assinatura digital ou um código, que o permite ler durante a semana toda, também sustenta um bocadinho as vendas do semanário. É mais um produto que estamos a dar aos leitores que são fiéis. Continuam a

compraram o Expresso, e também têm direito à oferta digital toda. Ou então fazem uma assinatura do Expresso Digital e têm acesso ao semanário em formato digital e ao Diário. Leem toda a oferta do Expresso por um custo menor.

**15. Relativamente aos sistemas que se usam para produzir notícias – Content station e o BackOffice. Que vantagens vê em implementar um sistema integrado de gestão no Expresso?**

As vantagens são todas, mas não tem sido possível. É muito mais fácil do que sair de uma plataforma e entrar noutra. Facilitaria coisas muito simples. Temos um texto fechado preparado para publicar no Diário, mas se houver um *site* a dar a notícia já não a podemos guardar, temos que a publicar no *site*. Porque é que temos que copiar tudo de uma plataforma para a outra? Não faz sentido nenhum. Mas tecnicamente não é possível com este sistema. O principal problema é a não integração do sistema, que permitisse gerir papel, diário e site na mesma plataforma. Se houvesse uma interface que me permitisse escolher onde publicar os artigos seria mais fácil. Quando republicamos as notícias – quando já perderam valor noticioso, novidade, atualidade – no *site*, temos que copiar tudo outra vez do *Content Station* para o *site*. Não faz sentido nenhum, mas deve ter a haver com questões financeiras, não sei. Ultrapassa-nos.

**E quanto ao Elvis?**

É um sistema de arquivo, não faz mais nada. Mas está bem feito, tem toneladas e toneladas de informação. Mas temos também um enorme arquivo físico. Quando precisamos que fotografias que só há em papel digitalizamos.

**16. Quais as principais utilidades do Google Analytics?**

Conseguimos saber exatamente o número de visualizações de cada notícia e quanto tempo uma pessoa esteve a ler em tempo real. E podemos formatá-lo para os vermos o que pretendemos. Todos os dias recebemos relatórios sobre a atividade das páginas.

**17. Qual a importância desses relatórios na tomada de decisão?**

É muito subjetivo. Eu não ligo muito a isso, porque há duas maneiras de ver os resultados. Se é o tema que está a ser mais lido, temos duas formas de ver: vou promove-lo mais nas redes sociais para aumentar ainda mais. Por outro lado, se perceber que um artigo importante não está a ser lido, posso promove-lo mais para as pessoas lerem. Podemos fazer as suas coisas. O primeiro principio é o bom senso. Se for uma notícia que é

‘espuma’, eu não vou puxar por ela, porque no Expresso não faz sentido. As notícias que têm morbidez ou sexo, ou coisas parecidas, têm mais potencial para as pessoas picarem e lerem. Mas acho que não é o campeonato do Expresso. Temos que dar formação boa, credível e que nos ensine qualquer coisa. Isso é muito difícil de gerir.

**18. Com essas estatísticas pode tentar perceber quais os temas mais lidos e criar novos padrões de leitura?**

Nós fazemos isso, usamos a informação das audiências para a informação futura. Notícias que tenham a ver com imobiliário, com consumo, sobretudo telecomunicações, energia e conselhos de poupanças são muito lidas. E, portanto, tentamos ter muitas e boas notícias boas nessas áreas. O difícil é nós darmos informação de outra maneira. Ou apresentá-la de outra maneira. Porque já praticamente tudo foi criado. Os jornais passam a vida a copiarem-se uns aos outros, quer nas temáticas, quer na apresentação dos temas. Porque já tudo foi feito. Mas há sempre novas formas de pegar nos temas. Isso pode dar muita coisa, pode dar uma noticiazinha, noticiazona ou uma manchete. Depende muito daquilo que se consegue, e de sorte. É preciso uma grande dose de sorte e outra de trabalho. É assim que se conseguem as grandes histórias. É uma questão de ir investigando, vai-se fazendo perguntas...

**19. Relativamente ao 2.59, quais foram as motivações que estão por detrás da criação desta rubrica?**

Foi a direção que a criou, penso que essencialmente tem a ver com a necessidade de dar informação de maneira diferente. Em segundo lugar, os tempos que correm e à velocidade a que a informação corre, à velocidade que as pessoas leem e não leem, dar muita informação, muitos dados, de uma forma diferente e em muito pouco tempo. O formato resultou, podia não ter resultado. A SIC já o copiou. É sobretudo isso, dar numa plataforma que é o digital, de um modo diferente e com grande rapidez para o leitor.

**Porque incidiram nos dados? Como é que as pessoas olham para os dados, para o jornalismo de dados?**

Eu acho que é importante. Porque se eu tiver em 3 minutos um maior volume possível de informação sobre os cursos superiores, por exemplo, facilitam-me muito a vida. Escuso de andar à procura no Google quais são os cursos com mais e menos saída, os valores das propinas, etc. Se eu tiver tudo sistematizado por alguém que teve esse trabalho, e se eu acredito que é uma fonte credível e com rigor, eu quero aquilo. O leitor vai ficar contente

porque tem em 3 minutos muita informação, que levaria horas a encontrar. O jornalismo de dados é cada vez mais importante porque toda a gente, todo o mundo recolhe dados. É cada vez uma forma mais importante de dar informação.

**20. Considera que uma notícia é mais credível aos olhos do público se conter dados?**

Pode tornar, se for pertinente. De certeza que os dados credibilizam ou podem credibilizar as notícias. Nem sempre há notícias que precisam de dados. Mas se tiverem dados credibiliza-as mais.

**21. Quais as competências que os jornalistas devem cultivar para que seja possível esse tratamento de dados?**

É preciso ter aprendizagem, alguma teoria, para saber ler um relatório, uma tabela, um gráfico. E preciso saber ler, mas não perder três dias a ler aquilo, mas apenas em uma ou duas horas. Saber como ler e depois é preciso ter algumas ferramentas como por exemplo dominar ferramentas técnicas informáticas para limpar a informação. Acho que a pessoa deve ter um curso avançado de *Excel*.

**22. Quanto às fontes utilizadas, há bases de dados contruídas pelo Expresso? Quais as vantagens?**

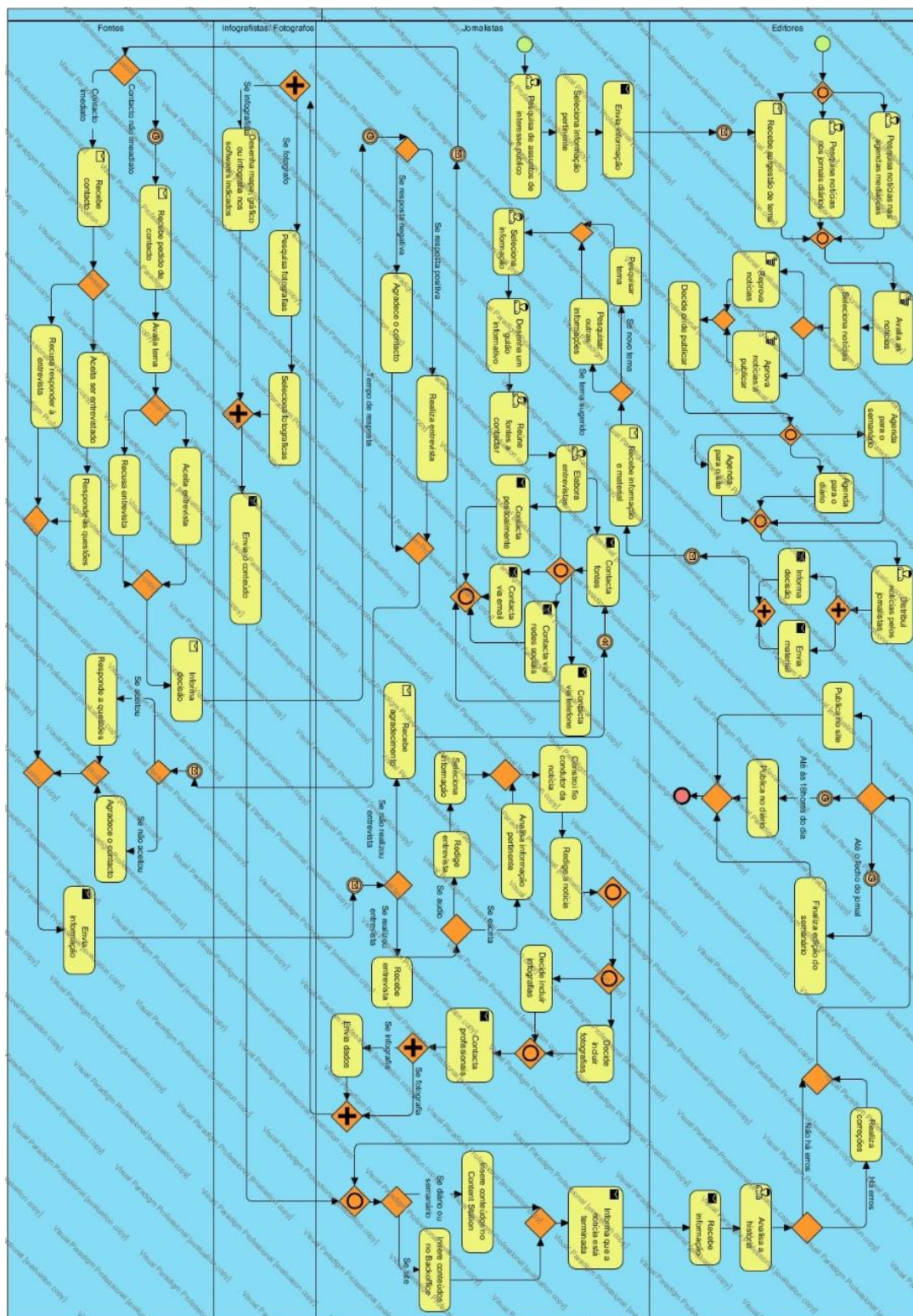
Nós temos informação e dados sobre determinados assuntos que são só nossos, porque as pessoas que tratam desses assuntos tratam de o fazer. Não temos bases de dados sistemáticas, é conforme os assuntos e conforme a disponibilidade e vontade de quem trata habitualmente dos assuntos de o fazerem. A infografia deve ter gráficos e tabelas sobre determinados assuntos que não devem haver sistematizados daquela maneira noutros sítios, porque vão recolhendo informação e vão construindo-os. Já temos publicado informação, um grande trabalho, há 4 anos, que tinha a ver com militares portugueses, que depois quiseram comprar, porque não tinham essa informação sistematizada dessa maneira. E ca esta outra saída económica, outra viabilização da informação, o que é acontece com a Gesco, que gere a informação do expresso e da SIC, gere o ELVIS, que fornece informação a outras organizações, que a pagam. Mas é outra fonte de rendimentos. Portanto tem dupla vantagem, criam algo único, extraem e dão aos leitores informação única, em primeira mão e depois ainda podem vende-la a outros, ganham dinheiro das duas maneiras.

**23. Quais as adaptações e as vantagens de introduzir o jornalismo de dados no jornal?**

Trabalhamos com dados quase diariamente, mas às vezes não os apresentamos em forma de dados. Mas vamos buscar a notícia e a informação aos dados. Vamos mergulhar nos dados, geralmente compilados por organismos nacionais e internacionais, como a fundação Francisco Manuel dos Santos, e vamos extrair dali informação. Às vezes a notícia é novidade, porque as pessoas sabem que saíram aqueles dados, mas não sabem que informação está lá dentro. Mas é necessário haver uma certa mudança cultural, sobretudo nos jornalistas mais antigos, porque muitas vezes a prosa contém muita informação que já está dita em gráficos e tabelas. Tenho lutado contra isso há anos e anos. A tabela é para simplificar, é para termos uma leitura mais direta. Se está na tabela não deve estar no texto, não deve haver repetições. Não me choca nada haver uma notícia com uma entrada e uma tabela, desde que a informação esteja lá. No digital, mapas e tabelas interativas é quase obrigatório todos os artigos terem, mas não é possível não há capacidade técnica, nem tempo.

## APÊNDICE 2. Modelação do processo de produção de notícias no jornal Expresso: desde a ideia inicial até à publicação

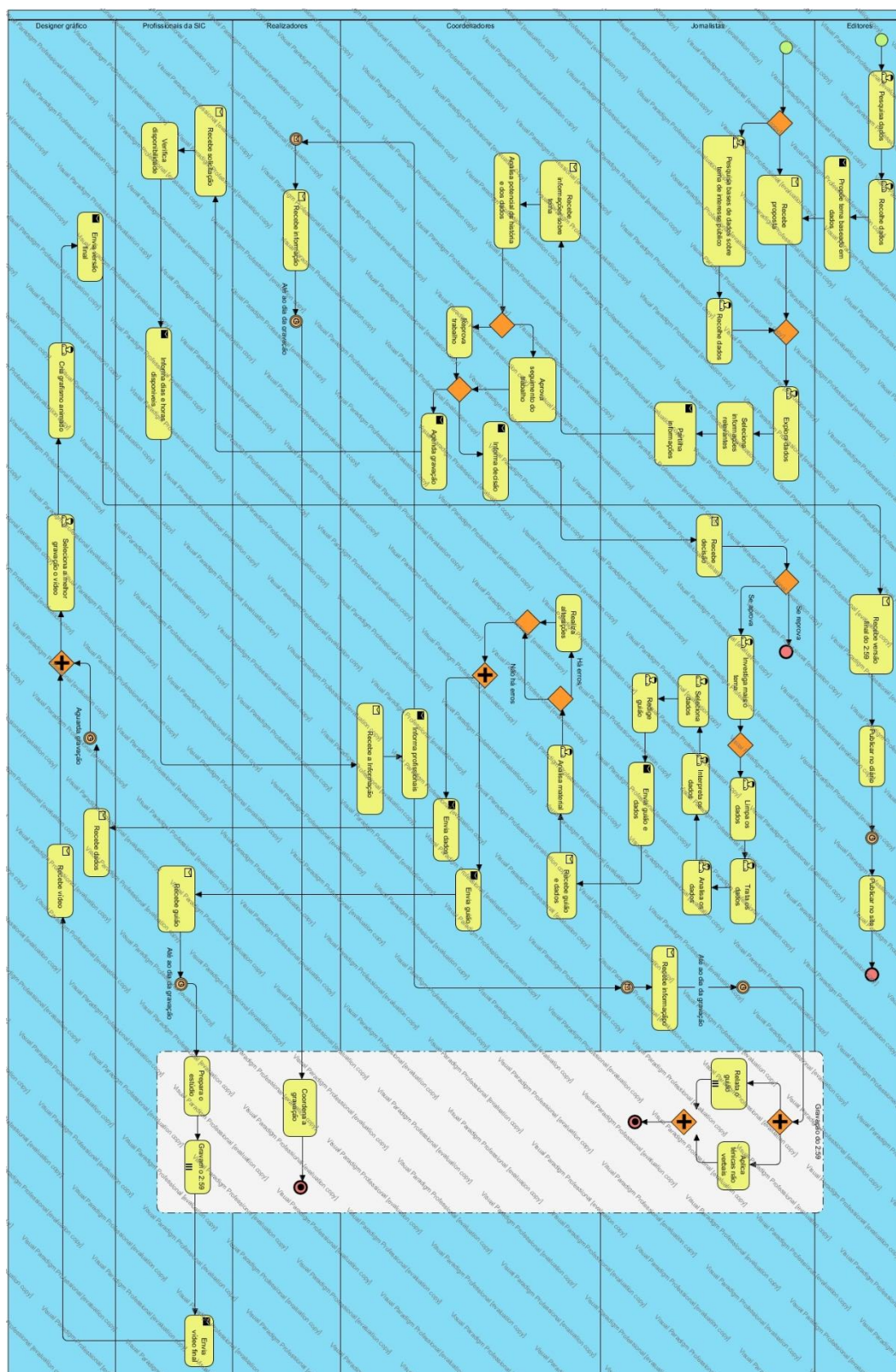
A presente modulação do processo de produção de notícias do jornal Expresso foi desenhada no *Visual Paradigm* 15.1, segundo a notação BPMN, e propõe-se a representar uma possível sequência do processo, sendo que é importante salientar que há nuances que podem alterar o seu rumo – algumas referidas no capítulo 3.





### APÊNDICE 3. Modelação do processo de produção da rubrica 2:59

A modulação do processo de produção da rubrica 2:59 procura representar as várias fases do processo e, em particular, os diferentes papéis desempenhados pelos seus intervenientes. A coreografia que se segue foi desenhada de acordo com a notação BPMN e com recurso ao *Visual Paradigm 15.1*.

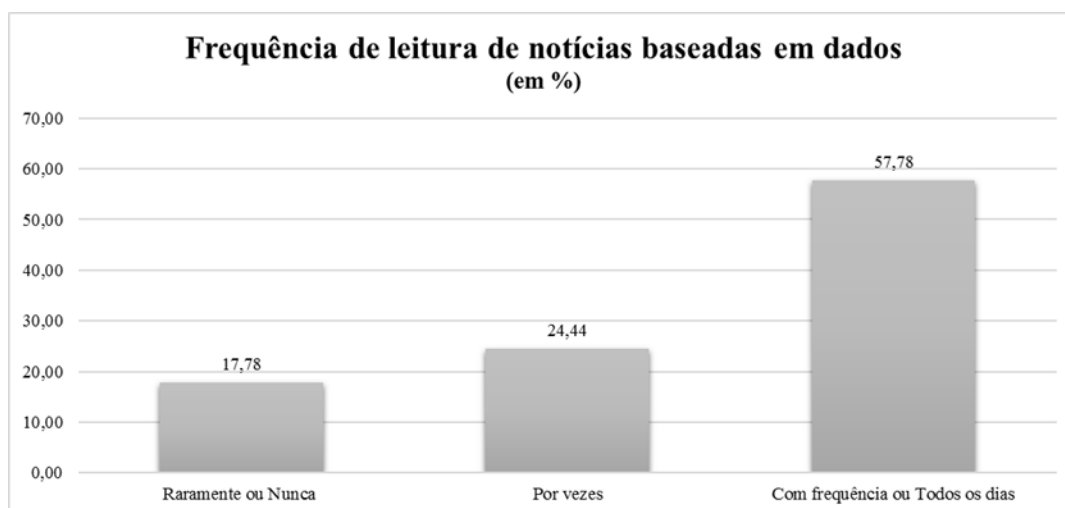


## **APÊNDICE 4. Visualizações referentes às percepções dos leitores quanto ao jornalismo de dados**

**Apêndice 4.1** - Aspetos que podem melhorar no jornalismo digital na perspetiva dos leitores assíduos de jornais digitais generalistas (em %)

<b>O que pode melhorar no jornalismo digital?</b>	<b>Discordo</b>	<b>Nem concordo nem discordo</b>	<b>Concordo</b>
Os conteúdos devem ter textos reduzidos	26,67	26,67	46,67
Os leitores devem decidir o percurso de leitura e não os jornalistas	31,11	33,33	35,56
Maior contextualização dos temas	2,22	13,33	84,44
É importante envolver mais os leitores no jornalismo recolhendo as suas opiniões	13,33	31,11	55,56
Investir em novas formas de interação com o conteúdo	0,00	13,33	86,67
Adaptar o conteúdo a diferentes plataformas ( <i>smartphones, ipads</i> )	4,44	8,89	86,67
Inserir mais conteúdos multimédia	4,44	17,78	77,78

**Apêndice 4.2** - Frequência de leitura de notícias baseadas em dados referente aos leitores que leem com frequência jornais digitais generalistas (em %)

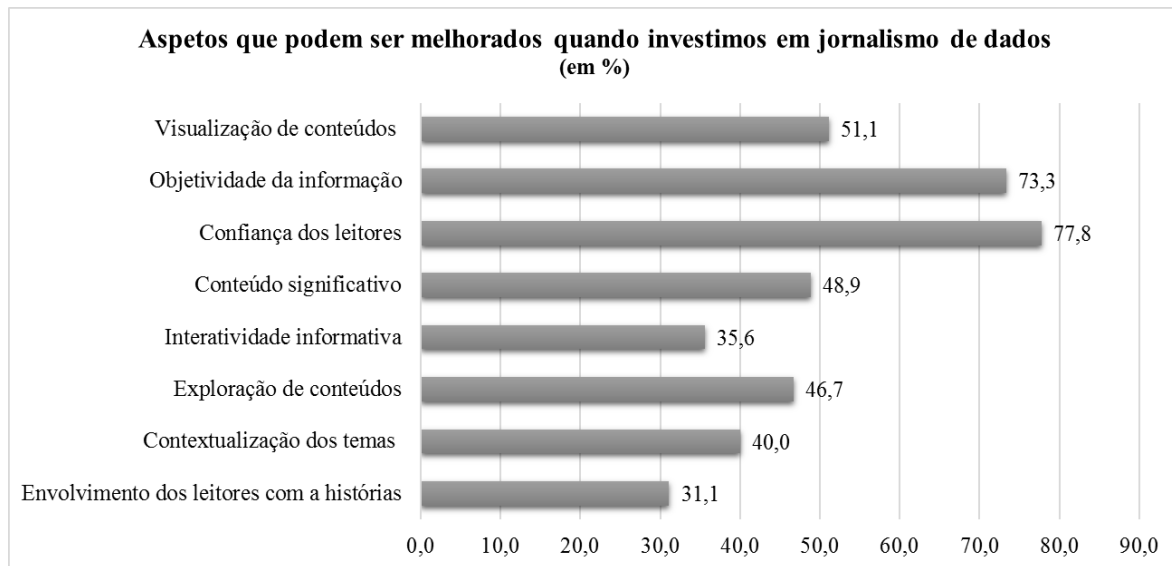




**Apêndice 4.3** – Percepção dos leitores, que leem com frequência jornais digitais generalistas, sobre as histórias baseadas em dados (em %)

<b>Percepção dos leitores sobre as histórias baseadas em dados</b>	<b>Discordo</b>	<b>Nem concordo nem discordo</b>	<b>Concordo</b>
Os dados tornam as notícias mais credíveis	0,0	11,1	88,9
Tenho uma maior percepção da realidade porque consigo quantificar ou qualificá-la	2,2	13,3	84,4
Os dados tornam a informação mais objetiva	8,9	15,6	75,6
Reconheço que a informação é única	6,7	46,7	46,7
Não presto atenção se a informação é ou não baseada em dados	71,1	17,8	11,1
Reconheço que é um trabalho diferenciado e atribuo-lhe valor	4,4	15,6	80,0
Considero que as histórias têm qualidade acrescida	6,7	22,2	71,1
Os conteúdos geralmente têm mais visualizações	4,4	44,4	51,1
Exploro o conteúdo para compreender melhor o significado dos dados	2,2	13,3	84,4
As histórias são mais complexas	24,4	31,1	44,4
Confio mais na informação quando contem dados	11,1	24,4	64,4
Geralmente os dados são extraídos de fontes fidedignas	8,9	42,2	48,9

**Apêndice 4.4** – Percepção dos leitores, que leem regularmente jornais digitais generalistas, quanto aos aspetos que podem melhorar quando se investe em jornalismo de dados (em %)



**APÊNDICE 5. *Short Paper* aceite na 13ª edição da Conferência Ibérica de Tecnologias e Sistemas de Informação (CISTI)**

# Um Olhar sobre o Jornalismo de Dados

## *An Insight into Data Journalism*

Vanessa Sofia Neves Abreu e Sousa

Instituto Superior de Contabilidade e Administração de Coimbra

Coimbra, Portugal

vanessa.abreusousa@outlook.pt

**Resumo** — É considerado uma nova tendência jornalística internacional, mas não é um processo novo, o jornalismo de dados — hoje suportado por tecnologias e sistemas de informação — procura desvendar informação relevante de uma avalanche de dados. A sua implementação gera mudanças: primeiro nas rotinas produtivas dos jornalistas e depois na experiência dos leitores. A partir da recolha das perceções dos leitores, através de questionários, e da sua análise, suportada pelo *Microsoft Office Excel 2016*, o presente artigo procura compreender o que promove uma aposta maior no jornalismo de dados.

**Palavras Chave** — *jornalismo de dados; sistemas de informação; big data; meio digital; leitores*

**Abstract** — *It is an international journalistic trend nowadays, but isn't a new process, the data journalism - today supported by information technologies and systems - is a way to find new information into a lot of data. The implementation of data journalism creates changes: first in productive routines of journalists and after in readers' experience. Based on the perceptions of readers collected, this article pretends to understand which promotes strong betting on data journalism.*

**Keywords** — *data journalism; information systems; big data; digital media; readers*

### Introdução

Descobrir informação relevante de uma imensidão de dados não é uma necessidade nova no jornalismo, mas perante uma era em praticamente tudo é passível de ser medido e quantificado, tornou-se mais importante do que nunca.

Aqui reside o grande destaque que o jornalismo de dados tem nos nossos dias: é este o fluxo de trabalho capaz de construir histórias que trazem algo de novo aos leitores, partindo da análise de uma quantidade substancial de dados [1]. São várias as habilidades e os saberes exigidos por este novo perfil emergente, o jornalista de dados. São várias também as mudanças que a integração desta prática nas redações acarreta. São igualmente várias — e até desconhecidas — as perceções que os leitores têm destas histórias baseadas em dados. As possibilidades que hoje são várias e abstratas, pretendem-se conhecer, para que seja possível torná-las concretas.

Neste sentido, este artigo pretende compreender e desvendar o mundo do jornalismo de dados, mostrando, numa primeira fase, quais são as perceções dos leitores relativamente às histórias baseadas em dados — as quais são apresentadas neste *shortpaper* — e numa segunda etapa, a realizar no futuro, quais são as capacidades desenvolvidas pelos jornalistas para desenvolver este tipo de conteúdos e quais as suas principais dificuldades.

### Contexto

#### *Jornalismo digital*

Se há poder que o jornalismo tem é o de se moldar às tendências tecnológicas que vão surgindo pelo mundo. Adaptou-se a novos meios de transmissão de informação — à rádio, à televisão e à Internet —, introduzindo novas técnicas, novas narrativas e novos modos de apresentação de conteúdos, os quais acabaram por caracterizar novas vertentes do jornalismo [8].

Da sua adaptação ao contexto digital surgem novas plataformas para aceder aos conteúdos jornalísticos: primeiro os *websites* informativos, e mais tarde as páginas nas redes sociais. O meio digital gera uma série de possibilidades no jornalismo: passa a ser possível interagir com os produtores das notícias assim como com outros leitores. A textualidade do jornalismo é renovada, introduzindo-se um novo conceito, o hipertexto que denota um texto em rede, no qual existem ligações a outros textos, permitindo assim que o leitor defina o seu próprio percurso de leitura. Também se torna possível integrar vários meios num só conteúdo como: texto, vídeo, som, gráficos, mapas, diagramas, entre outros [4]. Neste ambiente as limitações de espaço são quebradas, as histórias são contadas por vários meios e a comunicação entre jornalistas e leitores passa a ter dois sentidos.

O enraizar do ambiente digital no seio social e organizacional desencadeou a rápida geração de um grande volume de dados que são armazenados, o designado *Big Data*. Num momento em que tudo — ou quase tudo — pode ser quantificável e descrito através de dados, torna-se trivial extrair dos mesmos informação e sentido para compreender o mundo contemporâneo [5]. Aqui urge a importância atual do jornalismo de dados, pois permite descobrir valor nesses dados que de outra forma não seria perceptível [9].

#### *Jornalismo de dados: importância e desenvolvimento*

Tal como o jornalismo se redescobriu ao adaptar-se a novos meios de transmissão de informação, também as suas práticas se renovaram com a possibilidade de extrair informação relevante de um universo de dados com o auxílio das tecnologias mais recentes. A recolha, tratamento e análise de dados — um processo hoje suportado por sistemas e tecnologias de informação — com o propósito de reconhecer e contar histórias, não é um processo novo, mas hoje é considerado uma nova tendência internacional.

O jornalismo de dados já é um fluxo de trabalho inserido nas redações de vários meios de comunicação social, como por exemplo: *The New York Times*, *The Guardian*, *The Washington Post* e o jornal *Expresso*. O despertar para o seu valor e potencial surge na primeira década deste milénio, coincidindo com a investigação do caso *Wikileaks*. A necessidade de selecionar, analisar e compreender uma enorme quantidade de dados e informação existente em cerca de 92 mil documentos vazados do exército dos Estados Unidos — que revelaram a morte de milhares de civis na guerra do Afeganistão por mão de

militares norte-americanos –, evidenciou a importância do processo de jornalismo de dados na descoberta de histórias relevantes que de outra forma não seriam evidentes. Por este motivo, o caso *Wikileaks* é considerado uma força motriz do jornalismo de dados, revelando-se um processo capaz de encontrar e construir histórias complexas que nenhum outro processo jornalístico seria capaz de desvendar [2].

A necessidade de converter dados, quer quantitativos quer qualitativos, em informação relevante num contexto jornalístico é hoje maior do que alguma vez foi. É possível encontrar diversas fontes que disponibilizam dados, os quais traduzem realidades sobre várias temáticas de interesse público. Assim, tratá-los torna-se imprescindível para compreendê-los em profundidade [5].

Também em outros tempos os jornalistas sentiram necessidade de quantificar a realidade, de traduzi-la em números e compreendê-los. Na ausência de bases de dados públicas de organizações governamentais e não governamentais – que reportam a sua atividade periodicamente –, os jornalistas investigavam e recolhiam, eles próprios, os dados. Na ausência de sistemas e tecnologias de informação, que suportam o processo de tratamento e análise de dados, os jornalistas tentavam compreendê-los através de técnicas tradicionais: classificando, contando e cruzando-os manualmente. Uma das primeiras notícias baseadas em dados de que há registo foi publicada pelo jornal britânico *The Guardian* em 1821, início do século XIX, e procurava mostrar o “verdadeiro” número de alunos que recebiam educação gratuita nas escolas da cidade de Manchester [3].

A capacidade humana na recolha e tratamento de dados sem o auxílio de computadores é naturalmente limitada, pelo que o verdadeiro progresso desta forma de fazer jornalismo, de certo modo, coincide com o aparecimento de tecnologias e sistemas de informação, em meados dos anos 50. Surge, então, a denominada “Reportagem Assistida por Computador” que é considerada a primeira tentativa de explorar as potencialidades de um computador para recolher e analisar dados com intuito de produzir conteúdos jornalísticos. Mais tarde, nos anos 60, surge um novo conceito o “Jornalismo de precisão” definido por Philip Meyer, que propõe aproximar o jornalismo da ciência, introduzindo métodos científicos de investigação social no jornalismo, também com o apoio de tecnologias de informação. Propunha, então, capacitar os jornalistas em outras áreas para além das letras, introduzindo metodologias de recolha e tratamento de dados [3].

O jornalismo de dados – tal como o conhecemos hoje – resulta da evolução das vertentes jornalísticas acima referidas e pode ser sintetizado em quatro fases distintas. A primeira é a recolha de dados, na qual são conseguidas as bases de dados, quer criadas pelos repórteres a partir de documentos e relatórios técnicos, quer provenientes de bases de dados disponibilizadas por organizações governamentais e não-governamentais. Em segundo lugar surge a fase de tratamento, na qual decorre a seleção dos dados mais importantes para o estudo em questão, assim como a sua manipulação, através de sistemas de informação adequados. De seguida, destaca-se a fase de análise, caracterizada pela descoberta de tendências e anomalias e pela interpretação de resultados, promovendo a identificação de informações novas de interesse público. Por último é desenvolvido o produto final composto essencialmente por visualizações e um *design* informativo. A visualização de dados – através de vídeos, gráficos, quadros, mapas, diagramas e infografias – são baseadas nos resultados e destacam os aspetos mais relevantes, promovendo a sua compreensão e exploração. Já o *design* informativo permite enquadrar os dados num contexto, traduzindo-os em informação pertinente, que pode ser sustentada por declarações de especialistas na área e por outras histórias relacionadas. A visualização de dados combinada com um *design* informativo adequado facilita a compreensão das relações existentes entre os dados, bem como o seu contexto atual, promovendo, assim, o entendimento de histórias complexas por parte do público [6].

Para que todo este fluxo de trabalho seja exequível é importante capacitar os jornalistas de novos saberes, novas técnicas e métodos de trabalho, emergindo um novo perfil, o jornalista de dados [5]. A literacia estatística torna-se fundamental, bem como o uso ágil de ferramentas de recolha, processamento e visualização de dados, isto é, sistemas de informação. Além do referido, saber desenhar infografias, contruir gráficos, diagramas e programar aplicações são saberes que um jornalista de dados deve cultivar desde cedo. Genericamente, o perfil do jornalista de dados compila um conjunto de habilidades que vão além da arte de bem escrever, concentrando-se, essencialmente, no manuseamento das tecnologias e sistemas atuais, que suportam a gestão de bases de dados e a criação de visualizações [7].

A extração de informação de interesse público de um mar de dados para contar uma história tem um propósito e um destinatário: os jornalistas pretendem transmitir informação única e complexa de forma simplificada e facilmente compreensível pelo público. Estes conteúdos jornalísticos não só contam uma história, como também incluem – ou podem incluir – visualizações, como gráficos e infografias, que incorporam dados e informações, permitindo que os leitores explorem os conteúdos consoante as suas preferências pessoais. Se no papel estas visualizações são necessariamente estáticas, no meio digital podem ser desenhadas de forma interativa, abrindo caminho para um maior envolvimento dos leitores com o conteúdo apresentado. Assim, o público para além de ter a possibilidade de compreender informação complexa em profundidade, também pode explorar o conteúdo tendo por base os seus próprios interesses e motivações.

## Metodologia

O presente artigo foi desenvolvido a partir do suporte teórico – que foi selecionado de acordo com a pertinência do conteúdo – e da análise exploratória de dados, os quais foram recolhidos a partir de um questionário dirigido a todos os leitores de notícias.

O questionário foi elaborado na aplicação “Formulários Google” e o conjunto de questões realizadas são sobretudo fechadas. Para gerar a possibilidade de os inquiridos adicionarem aspetos diferentes dos apresentados também foi incluída uma questão semiaberta.

Relativamente às ferramentas de análise utilizadas, o *Microsoft Office Excel 2016* foi o sistema de informação utilizado para suportar, tratar, cruzar os dados e criar visualizações. No sentido de extrair informação dos dados foram realizados cálculos estatísticos – frequências absoluta e relativa – de forma a detetar tendências e anomalias nas opiniões dos leitores.

Embora um campo de investigação seja mais teórico e outro mais prático, os dois complementam-se na medida em que a revisão de literatura permite aprofundar os conhecimentos sobre a temática, tornando a construção do questionário mais precisa e próxima do tema em estudo e a interpretação de resultados mais sustentada.

## Seleção de Dados

Os dados recolhidos são fruto de um questionário desenhado tendo por base a necessidade de compreender as perceções dos leitores quando exploram conteúdos digitais baseados em dados, segundo um pressuposto da investigadora fundamentado na literatura.

O questionário foi disponibilizado nas redes sociais, nomeadamente no *Facebook*, entre 22 de janeiro e 5 de fevereiro de 2018 e a análise seguinte tem por base uma amostra de 91 respostas recolhidas nesse período. Já o universo do estudo passa por todos os consumidores de notícias num contexto digital.

Divido em três seções, o questionário pretendeu recolher os dados sociodemográficos dos respondentes e também outras informações que permitem traçar os seus perfis enquanto leitores. A última seção é destinada a compreender quais as perceções existentes quanto aos

conteúdos jornalísticos baseados em dados, e o que pode melhorar no jornalismo digital com uma maior integração do fluxo de trabalho em estudo nas redações.

## Objetivos

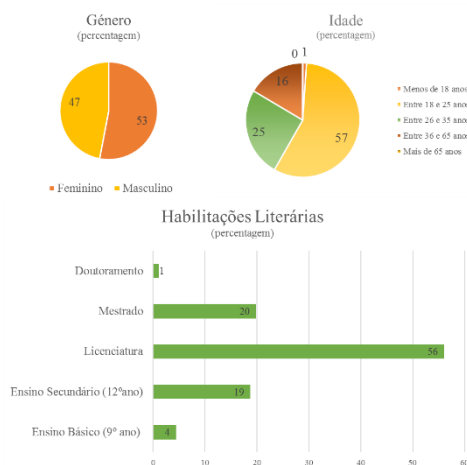
O presente estudo procura compreender de que forma as histórias baseadas em dados são entendidas pelo público e de que forma são consultadas. Pretende-se também decifrar quais são as plataformas digitais em que a sua integração fará mais sentido, partindo da análise de frequência de consulta de notícias nas mesmas. E por último, cruzar as principais lacunas apontadas pelos leitores no jornalismo digital, com os principais aspetos que o jornalismo de dados pode promover será também uma meta a atingir.

Numa próxima fase pretende-se completar o estudo com a visão dos jornalistas, dirigindo-lhes um outro questionário. Neste sentido será possível compreender também quais as principais tarefas que fazem parte do fluxo de trabalho dos jornalistas quando trabalham com dados e quais são executadas com mais e menos dificuldade. Num segundo ponto, também se pretende descobrir quais são as principais fontes e sistemas de informação usados pelos jornalistas, bem como a frequência de produção de conteúdos baseados em dados.

## Resultados

### Caracterização sociodemográfica da amostra

Os resultados da primeira secção do questionário, representados no *Dashboard 1*, permitem traçar um perfil sociodemográfico genérico com base nas tendências: há mais mulheres do que homens a responder ao questionário – embora a diferença seja pouco expressiva de apenas 5 pessoas –, a maior parte dos inquiridos tem idades entre os 18 e os 25 anos e completaram a licenciatura.



Dashboard 1. Caracterização sociodemográfica dos inquiridos

### Identificação do Perfil de Leitores

Um ponto a compreender na segunda secção do questionário traduz-se na seguinte questão: como as notícias chegam até ao público? Verificou-se que os inquiridos tanto tomam iniciativa de consultar os meios de comunicação com frequência, como também leem muitas vezes os conteúdos jornalísticos que surgem das redes sociais. Ainda assim, há ligeiramente mais pessoas a ler as notícias que surgem nas redes sociais (49%) do que as que assumem tomar iniciativa de consultar os meios de comunicação (45%).

Quanto às plataformas digitais mais acedidas destacam-se os jornais generalistas, em que 49% dos inquiridos dizem consultar com frequência ou todos os dias. Também os jornais exclusivamente online têm uma frequência de acesso pelos leitores inquiridos considerável: 42% leem-nos com frequência ou todos os dias. As plataformas digitais

televisivas, radiofónicas, bem como os jornais regionais e revistas em formato digital são as menos consultadas pela amostra.

### O jornalismo de dados no meio digital

Um dos pontos abordados na terceira secção do questionário procura compreender quais os pontos do jornalismo digital que podem ser melhorados na perspetiva dos leitores. Dos sete aspetos mencionados cinco revelaram uma concordância genérica por maior parte da amostra, tal como é possível verificar na figura 1, nomeadamente: a maior contextualização dos temas, um envolvimento maior dos leitores no jornalismo recolhendo as suas opiniões, um investimento acentuado em novas formas de interação com o conteúdo, uma maior adaptação do conteúdo a diferentes plataformas e introdução de conteúdos multimédia.

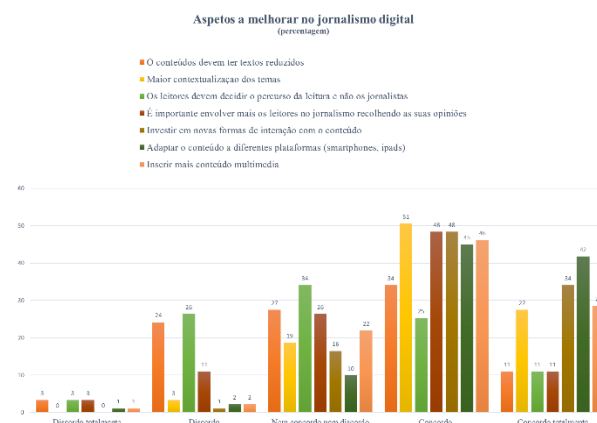


Figura 1. Aspetos a melhorar no jornalismo digital

Num segundo ponto foi possível averiguar que precisamente metade dos inquiridos lê notícias baseadas em dados com frequência ou todos os dias, como demonstra a figura 2, enquanto mais de um terço da amostra também admite fazê-lo, embora com uma frequência menor – apenas por vezes.

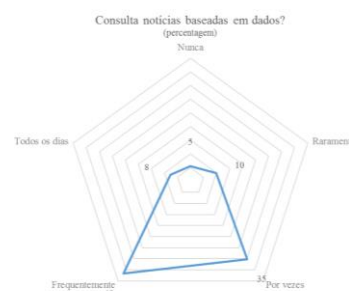


Figura 2. Frequência de leitura notícias baseadas em dados

A reflexão por parte dos leitores quanto às notícias baseadas em dados também gerou consenso em grande fatia dos aspetos referidos, como revela a figura 3. A maior parte dos leitores considera que estes conteúdos promovem uma maior perceção da realidade pois é possível quantificar ou qualificá-la e a informação que os constitui é mais credível e objetiva. Também a maioria considera que estas histórias são mais complexas, têm qualidade acrescida e valor por considerar que é um trabalho diferenciado.

No geral, os leitores dizem confiar mais nas histórias jornalísticas quando apresentam dados, assim como assumem explorar melhor os conteúdos para compreender o significado dos dados. Exatamente metade dos inquiridos concorda que as histórias baseadas em dados têm

geralmente mais visualizações, mas 47% dos respondentes assumem não ter opinião sobre a questão. Também há mais leitores a não ter uma opinião definida quanto à utilização de fontes fidedignas na recolha dos dados, do que aqueles que concordam com a afirmação.

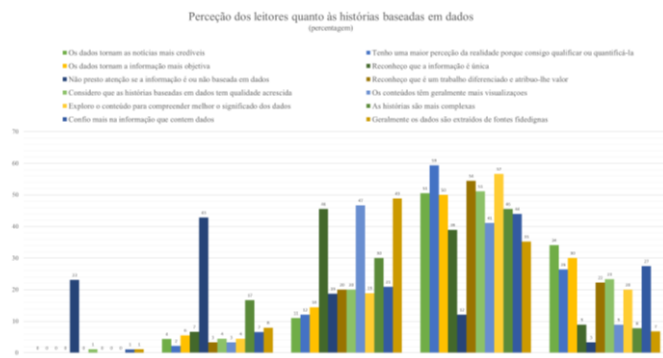


Figura 3. Percepção dos leitores quanto às notícias baseadas em dados

A última questão do questionário – O que promove um maior investimento em jornalismo de dados? – vem descobrir quais as necessidades dos leitores que o jornalismo de dados no seio digital pode preencher. O que mais promove esta vertente do jornalismo é a confiança dos leitores – aspeto nomeado por 66 inquiridos – como representa a figura 4. Também a maioria diz que as histórias baseadas em dados promovem a objetividade da informação e a criação de conteúdo significativo. Também com nomeações expressivas está a promoção de visualizações de conteúdos (vídeo, gráficos e infografias), a maior exploração de conteúdos e envolvimento dos leitores com a história. Em resposta aberta, os inquiridos acrescentam que um investimento maior neste fluxo de trabalho melhora também a assertividade jornalística, facilita a leitura dos conteúdos e promove também a solução de problemas.



Figura 4. Aspetos que podem melhorar com um maior investimento em jornalismo de dados

## Conclusão

As histórias baseadas em dados não passam despercebidas pelo público, são aliás lidas com uma frequência notável. A confiança no jornalismo – que hoje se encontra em crise – pode ser restaurada com um maior investimento nestes conteúdos, que aos olhos do público são credíveis e objetivos. Os leitores assumem que os dados inseridos nas histórias não só permitem quantificar e qualificar uma realidade, como também explorar o conteúdo para melhor compreender o seu significado. Um maior valor e qualidade são dois benefícios destacados pelo público em estudo, que caracterizam os conteúdos baseados em dados.

Estes pontos positivos evidenciados pela amostra revelam que investir em jornalismo de dados é uma mais valia para os meios de comunicação, na medida em que os conteúdos – resultantes desse processo – são valorizados pelo público e transmitem maior confiança. No meio digital, os conteúdos jornalísticos são consultados – ou por iniciativa própria dos leitores ou porque surgem nas redes sociais – sobretudo em jornais generalistas e jornais exclusivamente online, pelo que uma maior integração do jornalismo de dados poderá ter maior sucesso nos meios de comunicação deste tipo. Por outro lado, um investimento nesta vertente jornalística poderá atrair novos leitores para os jornais regionais e revistas digitais, por exemplo.

Foram vários os pontos que geraram consenso quanto aos aspetos a serem melhorados no jornalismo digital. Também foram vários os aspetos nomeados que indicam o que pode melhorar com um maior investimento em jornalismo de dados. Do cruzamento das lacunas do jornalismo digital com as vantagens do jornalismo de dados há três pontos que coincidem. As histórias baseadas em dados promovem tanto a exploração de conteúdos como o envolvimento dos leitores com a história, aproximando assim os leitores do jornalismo – um vazio apontado pela maioria da amostra. A integração de mais visualizações, como gráficos e infografias, nestas histórias colmata também a necessidade dos leitores de explorar conteúdos multimédia. Se estas visualizações forem interativas – possibilidade exclusiva do meio digital –, a necessidade dos leitores de interagir com os conteúdos jornalísticos fica também preenchida.

Posto isto, investir em jornalismo de dados pode satisfazer as necessidades sentidas pelos leitores, restaurando, ao mesmo tempo, a sua confiança no jornalismo.

## Agradecimentos

Quero agradecer às professoras Isabel Pedrosa e Sara Meireles pela motivação e esclarecimentos e aos editores José Cardoso e Germano Oliveira do jornal “Expresso” pela clarificação da temática em estudo. A todos os que responderam ao questionário, o meu muito obrigada.

## Referências Bibliografia

- Affelt, A. (s.d.). Investigative Journalism and Big Data. The accidental data scientist.
- Baack, S. (2011). A new style of news reporting : Wikileaks and Data-Driven Journalism. SSOAR.
- Bounegru, L. (2012). The Data Journalism Handbook. Retrieved July 15, 2017, from [http://datajournalismhandbook.org/pt/introducao\\_4.html](http://datajournalismhandbook.org/pt/introducao_4.html)
- Canavilhas, J. (s.d.). Webjornalismo: da pirâmide invertida à pirâmide deitada. Universidade da Beira Interior.
- Góis, D., Marques, R., & Brunet, K. S. (2013). Jornalismo de Dados e Mapping Art: Possíveis aproximações. In 8º SOPCOM (pp. 49–55)
- Henninger, M. (2013). Data-Driven Journalism. Em A. Knight, CHALLENGE AND CHANGE: Reassessing Journalism's Global Future (pp. 158-183). UTS ePRESS.
- López-García, X., Toural, C., & Rodríguez-Vázquez, A.-I. (2016). Software, estadística y gestión de bases de datos en el perfil del periodista de datos. El Profesional de La Información, 25(2), 286. <https://doi.org/10.3145/epi.2016.mar.16>
- Oliveira, S., & Torres, B. V. (2013). O paradigma “Jornalismo Digital em Base de Dados”: modos de narrar, formatos e visualização para conteúdos. Galaxia, 25, 152–164.
- Teixeira, W., & Junior, L. (2012). Big Data, Jornalismo Computacional e Data Journalism: estrutura, pensamento e prática profissional na Web de dados. Estudo Em Comunicação nº12, 207–222

